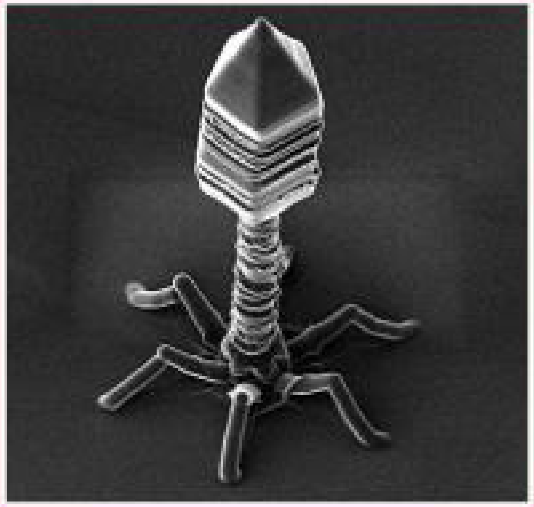
Woodhead Publishing Series in Electronic and Optical Materials



Nanolithography

The art of fabricating nanoelectronic and nanophotonic devices and systems

Edited by Martin Feldman

WP
WOODHEAD PUBLISHING

Nanolithography

Related titles:

Carbon nanotubes and graphene for photonic applications
(ISBN 978-0-85709-417-9)

Laser growth and processing of photonic devices
(ISBN 978-1-84569-936-9)

Handbook of solid-state lasers
(ISBN 978-0-85709-272-4)

Details of these books and a complete list of titles from Woodhead Publishing can be obtained by:

- visiting our web site at www.woodheadpublishing.com
- contacting Customer Services (e-mail: sales@woodheadpublishing.com; fax: +44 (0) 1223 832819; tel.: +44 (0) 1223 499140 ext. 130; address: Woodhead Publishing Limited, 80, High Street, Sawston, Cambridge CB22 3HJ, UK)
- in North America, contacting our US office (e-mail: usmarketing@woodheadpublishing.com; tel.: (215) 928 9112; address: Woodhead Publishing, 1518 Walnut Street, Suite 1100, Philadelphia, PA 19102-3406, USA)

If you would like e-versions of our content, please visit our online platform: www.woodheadpublishingonline.com. Please recommend it to your librarian so that everyone in your institution can benefit from the wealth of content on the site.

We are always happy to receive suggestions for new books from potential editors. To enquire about contributing to our Electronic and Optical Materials series, please send your name, contact address and details of the topic/s you are interested in to laura.pugh@woodheadpublishing.com. We look forward to hearing from you.

The team responsible for publishing this book:

Commissioning Editor: Laura Pugh
Publications Coordinator: Lucy Beg
Project Editor: Elizabeth Moss
Editorial and Production Manager: Mary Campbell
Production Editor: Richard Fairclough
Project Manager: Newgen Knowledge Works Pvt Ltd
Copyeditor: Newgen Knowledge Works Pvt Ltd
Proofreader: Newgen Knowledge Works Pvt Ltd
Cover Designer: Terry Callanan

Woodhead Publishing Series in Electronic and Optical Materials:
Number 42

Nanolithography

The art of fabricating nanoelectronic and nanophotonic devices and systems

---

Edited by
Martin Feldman

WOODHEAD PUBLISHING

Oxford Cambridge Philadelphia New Delhi

Published by Woodhead Publishing Limited,
80 High Street, Sawston, Cambridge CB22 3HJ, UK
www.woodheadpublishing.com
www.woodheadpublishingonline.com

Woodhead Publishing, 1518 Walnut Street, Suite 1100, Philadelphia,
PA 19102-3406, USA

Woodhead Publishing India Private Limited, 303 Vardaan House, 7/28 Ansari Road,
Daryaganj, New Delhi – 110002, India
www.woodheadpublishingindia.com

First published 2014, Woodhead Publishing Limited
© Woodhead Publishing Limited, 2014, except Chapter 6 which is © H. Smith, 2014. The publisher has made every effort to ensure that permission for copyright material has been obtained by authors wishing to use such material. The authors and the publisher will be glad to hear from any copyright holder it has not been possible to contact. The authors have asserted their moral rights.

This book contains information obtained from authentic and highly regarded sources. Reprinted material is quoted with permission, and sources are indicated. Reasonable efforts have been made to publish reliable data and information, but the authors and the publisher cannot assume responsibility for the validity of all materials. Neither the authors nor the publisher, nor anyone else associated with this publication, shall be liable for any loss, damage or liability directly or indirectly caused or alleged to be caused by this book.

Neither this book nor any part may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, microfilming and recording, or by any information storage or retrieval system, without permission in writing from Woodhead Publishing Limited.

The consent of Woodhead Publishing Limited does not extend to copying for general distribution, for promotion, for creating new works, or for resale. Specific permission must be obtained in writing from Woodhead Publishing Limited for such copying.

Trademark notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation, without intent to infringe.

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library.

Library of Congress Control Number: 20131948203

ISBN 978-0-85709-500-8 (print)
ISBN 978-0-85709-875-7 (online)
ISSN 2050-1501Woodhead Publishing Series in Electronic and Optical Materials (print)
ISSN 2050-151X Woodhead Publishing Series in Electronic and Optical Materials (online)

The publisher's policy is to use permanent paper from mills that operate a sustainable forestry policy, and which has been manufactured from pulp which is processed using acid-free and elemental chlorine-free practices. Furthermore, the publisher ensures that the text paper and cover board used have met acceptable environmental accreditation standards.

Typeset by Newgen Knowledge Works Pvt Ltd, India
Printed by Lightning Source

Contents

Contributor contact details *xiii*
Woodhead Publishing Series in Electronic and Optical Materials *xvii*
Preface *xxiii*

1 Optical projection lithography 1
B. W. Smith, Rochester Institute of Technology, USA

1.1 Introduction 1
1.2 Lithography technology and trends 1
1.3 Fundamentals of optical lithography 5
1.4 Image evaluation 8
1.5 Projection lithography systems 13
1.6 Wavelengths for optical lithography 22
1.7 Lithography in the deep ultraviolet (UV) 23
1.8 Resolution enhancement technology 27
1.9 Immersion lithography 36
1.10 Multiple patterning optical lithography 38
1.11 Conclusion 40
1.12 References 40

2 Extreme ultraviolet (EUV) lithography 42
B. J. Rice, SEMATECH, USA

2.1 Introduction 42
2.2 EUV sources 52
2.3 EUV optics 58
2.4 EUV masks 62
2.5 EUV resists 70
2.6 EUV integration and implementation challenges 73
2.7 Conclusion and future trends 75
2.8 Acknowledgments 76
2.9 References 77

3 Electron beam lithography 80
T. R. Groves, University at Albany (SUNY), USA

3.1 Introduction 80
3.2 Using pixel parallelism to address the throughput bottleneck 84
3.3 The tradeoff between resolution and throughput 96
3.4 Distributed systems 100
3.5 Ultimate lithographic resolution 104
3.6 Electron-beam patterning of photomasks for optical lithography 107
3.7 Conclusion 110
3.8 Acknowledgements 111
3.9 References 111

4 Focused ion beams for nano-machining and imaging 116
M. Utlaut, University of Portland, USA

4.1 Introduction 116
4.2 An adumbrated history of focused ion beams (FIBs) 118
4.3 Sources of ions: a quartet of types 119
4.4 Charged particle optics 125
4.5 Ion-matter interactions 128
4.6 Milling 138
4.7 Deposition 145
4.8 Imaging 147
4.9 Spectroscopy 153
4.10 Conclusion and future trends 153
4.11 References 155

5 Masks for micro- and nanolithography 158
E. Gallagher and M. Hibbs, IBM Microelectronics Inc., USA

5.1 Introduction 158
5.2 Mask materials 162
5.3 Mask process 167
5.4 Mask metrology 168
5.5 Defects and masks 172
5.6 Conclusion 177
5.7 References 178

6 Maskless photolithography 179
M. E. Walsh and F. Zhang, LumArray, Inc., USA,
R. Menon, University of Utah, USA, and H. I. Smith,
LumArray, Inc. and Massachusetts Institute of
Technology, USA

6.1 Introduction 179
6.2 The use of photons as opposed to charged particles 180
6.3 Forms of maskless photolithography 181
6.4 Zone-plate-array lithography (ZPAL) 183
6.5 Proximity-effect correction 186
6.6 Extending the resolution of ZPAL 187
6.7 Commercialization of ZPAL by LumArray, Inc. 189
6.8 Conclusion 191
6.9 References 192

7 Chemistry and processing of resists for nanolithography 194
A. Novembre, Princeton University, USA and
S. Liu, IBM Corporation, USA

7.1 Introduction 194
7.2 Resists for optical lithography: synthesis and radiation induced chemistry of resists as a function of exposure technology 200
7.3 Chemically amplified resist process considerations 215
7.4 Chemically amplified resists for 193 nm lithography 217
7.5 Resists for extreme ultraviolet lithography (EUVL) 231
7.6 Resists for electron beam lithography 234
7.7 Resists for selected forward looking lithographic technologies 265
7.8 Resist resolution limitations 268
7.9 Conclusion 270
7.10 References 271

8 Directed assembly nanolithography 287
S. Matsui, University of Hyogo, Japan, M. Takenaka,
Kyoto University, Japan and H. Yoshida, Hitachi Research
Laboratory, Japan

8.1 Introduction 287
8.2 Block copolymers in lithography 287
8.3 Directed self-assembly of block copolymers 294

8.4 Programmable three-dimensional lithography 302
8.5 Conclusion 310
8.6 References 311

9 Nanoimprint lithography 315
D. Resnick, Molecular Imprints Inc., USA

9.1 Introduction 315
9.2 An overview of imprint lithography 320
9.3 Soft lithography 321
9.4 Thermal imprint lithography 322
9.5 Alternative thermal imprint processes 327
9.6 Ultraviolet (UV) nanoimprint lithography overview 330
9.7 Jet and flash imprint lithography 331
9.8 Roll to roll imprint lithography 340
9.9 Defectivity 342
9.10 Conclusions 344
9.11 Acknowledgments 345
9.12 References 345

10 Nanostructures: fabrication and applications 348
X. Cheng, Texas A&M University, USA

10.1 Introduction 348
10.2 Characterization of nanostructures 350
10.3 Methods to create nanostructures: top-down fabrication of nanostructures 350
10.4 Methods to create nanostructures: bottom-up fabrication of nanostructures 358
10.5 Properties of nanostructures 364
10.6 Applications of nanostructures 370
10.7 References 371

11 Nanophotonics: devices for manipulating light at the nanoscale 376
P. Dastmalchi, A. Haddadpour and G. Veronis, Louisiana State University, USA

11.1 Introduction 376
11.2 Photonic crystals 377
11.3 Ring resonators 379
11.4 Extraordinary optical transmission through subwavelength apertures 382

11.5 Optical nanoantennas 384
11.6 Plasmonic focusing 387
11.7 Near-field optical microscopy 390
11.8 Plasmonic waveguides 392
11.9 Enhancement of nonlinear processes 393
11.10 Application in photovoltaics 395
11.11 Conclusion 395
11.12 References 396

12 Nanodevices: fabrication, prospects for low dimensional devices and applications 399
T. Daniels-Race, Louisiana State University, USA

12.1 Introduction 399
12.2 Motivation for nanodevices 401
12.3 Nanofabrication: creating the building blocks for devices 403
12.4 Prospects for low dimensional devices 408
12.5 Beyond the bottom-up: hybrid nanoelectronics 413
12.6 Conclusion and future trends 417
12.7 References 417

13 Microfluidics: technologies and applications 424
L. Jiang and N. S. Korivi, Tuskegee University, USA

13.1 Introduction 424
13.2 Current trends in microfluidics 425
13.3 Present state of technology 429
13.4 Applications 434
13.5 Future trends 436
13.6 Conclusion 440
13.7 Sources of further information and advice 440
13.8 References 441

14 Modeling of nanolithography processes 444
A. Isoyan and L. S. Melvin III, Synopsys Inc., USA

14.1 Introduction 444
14.2 Optical lithography modeling 445
14.3 The optical system in optical lithography modeling 449
14.4 Photoresist model 453
14.5 Model critical dimension (CD) extraction 454
14.6 Difficulties in modeling 455

14.7 Extreme ultraviolet (EUV)/electron beam lithography modeling 457
14.8 Conclusion 462
14.9 References 462

15 Mask-substrate alignment via interferometric moiré fringes 466
E. E. Moon, Massachussetts Institute of Technology, USA

15.1 Introduction 466
15.2 Background to alignment methods 467
15.3 Fundamentals of interferometric-spatial-phase imaging 472
15.4 Implementation of moiré 475
15.5 Characteristics of moiré fringe formation 479
15.6 Performance of ISPI 494
15.7 Backside ISPI 497
15.8 Conclusion and future trends 500
15.9 References 501

16 Sidewall roughness in nanolithography: origins, metrology and device effects 503
V. Constantoudis and E. Gogolides, NCSR Demokritos, Greece and G. P. Patsis, Technological Educational Institution of Athens, Greece

16.1 Introduction 503
16.2 Metrology and characterization 505
16.3 Process and material effects: modeling and simulation 513
16.4 Process and material effects: experimental results 515
16.5 Impact on device performance 523
16.6 Conclusions 530
16.7 References 532

17 New applications and emerging technologies in nanolithography 538
F. Yesilkoy, C. Ropp, Z. Cummins, R. Probst, E. Waks, B. Shapiro and M. Peckerar, University of Maryland, USA

17.1 Introduction 538
17.2 Applications of high-resolution patterning to new device structures: advances in tunneling structures 541
17.3 Geometry control of the tunnel junctions 543
17.4 The quantum dot placement problem 547

17.5 Conclusion 548
17.6 Acknowledgments 549
17.7 References 549

Index *551*

Contributor contact details

(* = main contact)

Editor

Martin Feldman
Division of Electrical and Computer Engineering
Patrick Taylor Hall
Louisiana State University
Baton Rouge
LA 70803, USA

E-mail: mfeldm1@lsu.edu

Chapter 1

Bruce W. Smith
Microsystems Engineering
College of Engineering
Rochester Institute of Technology
168 Lomb Memorial Drive
Rochester
NY 14623, USA

E-mail: bwsemc@rit.edu

Chapter 2

Bryan J. Rice
Intel Corporation assignee to SEMATECH
257 Fuller Rd, Suite 2200
Albany
NY 12203, USA

E-mail: bryan.rice@sematech.org

Chapter 3

Timothy R. Groves
College of Nanoscale Science and Engineering
University at Albany, State University of New York
257 Fuller Road
Albany
NY 12203, USA

E-mail: tgroves@albany.edu

Chapter 4

Mark Utlaut
Department of Physics
University of Portland
MSC121
5000 N. Willamette Blvd.
Portland
Oregon 97203, USA

E-mail: utlaut@up.edu

Chapter 5

Emily Gallagher* and Michael Hibbs
IBM Microelectronics Inc.
1000 River Street
Essex Junction
VT 05452, USA

E-mail: fisch@us.ibm.com and mhibbs@us.ibm.com

Chapter 6

Michael Walsh and
Feng Zhang
LumArray, Inc.
15 Ward Street
Somerville
MA 02143, USA

Rajesh Menon
University of Utah
Salt Lake City
UT
USA

Henry I. Smith*
LumArray, Inc.
15 Ward Street
Somerville
MA 02143, USA

and

Massachusetts Institute of Technology
Cambridge
MA 02139, USA

E-mail: hismith@mit.edu

Chapter 7

Anthony Novembre*
Princeton Institute for the Science and Technology of Materials
Princeton University
70 Prospect Ave.
Princeton
NJ 08540, USA

E-mail: novembre@Princeton.EDU

Sen Liu
IBM Corporation
2070 Route 52
Hopewell Junction
NY 12533, USA

E-mail: liuse@us.ibm.com

Chapter 8

Shinji Matsui*
Laboratory of Advanced Science and Technology
University of Hyogo
3–1-2 Koto, Kamigori
Ako
Hyogo 678–1205, Japan

E-mail: matsui@lasti.u-hyogo.ac.jp

Mikihito Takenaka
Department of Polymer Chemistry, Graduate School of Engineering
Kyoto University
Kyotodaigaku-katsura
Nishikyo-ku
Kyoto 615–8510, Japan

E-mail: takenaka@alloy.polym.kyoto-u.ac.jp

Hiroshi Yoshida
Department of Organic Materials Research
Hitachi Research Laboratory, Hitachi, Ltd.
7-1-1 Omika
Hitachi-city
Ibaraki, 319–1292, Japan

E-mail: hiroshi.yoshida.jz@hitachi.com

Chapter 9

Doug Resnick
Molecular Imprints Inc.
1807 West Braker Lane
Austin, TX USA 78758 USA

E-mail: DResnick@molecularimprints.com

Chapter 10

Xing Cheng
Department of Electrical and Computer Engineering
Texas A&M University
College Station
Texas 77843, USA

E-mail: chengx@ece.tamu.edu

Chapter 11

Pouya Dastmalchi, Ali Haddadpour and Georgios Veronis*
School of Electrical Engineering & Computer Science and Center for Computation & Technology
Louisiana State University
3101 Patrick F. Taylor Hall Baton Rouge Louisiana 70803, USA

E-mail: gveronis@lsu.edu, gveronis@gmail.com

Chapter 12

Theda Daniels-Race
Division of Electrical and Computer Engineering
School of Electrical Engineering and Computer Science
Louisiana State University
3197 Patrick F. Taylor Hall
Baton Rouge
LA 70803, USA

E-mail: tdrace@lsu.edu

Chapter 13

Li Jiang and Naga S. Korivi
Electrical Engineering Department
Tuskegee University
Tuskegee,
Alabama, 36088, USA

E-mail: ljiang@mytu.tuskegee.edu

Chapter 14

Artak Isoyan and Lawrence S. Melvin* III
Synopsys Inc. 2025 NW Cornelius Pass Road
Hillsboro
OR 97124, USA

E-mail: Isoyan@synopsys.com and lmelvin@synopsys.com

Chapter 15

Euclid E. Moon
Massachusetts Institute of Technology
Cambridge
MA 02139, USA

E-mail: euclid@nano.mit.edu

Chapter 16

Vassilios Constantoudis* and Evangelos Gogolides
Institute of Microelectronics
NCSR 'Demokritos'
Patr. Grigoriou and Neapoleos str.
Aghia Paraskevi 15310, Greece

E-mails: vconst@imel.demokritos.gr evgog@imel.demokritos.gr

George P. Patsis
Department of Electronics
Technological Educational Institution of Athens
Aegaleo 12210, Greece

E-mail: patsisg@teiath.gr

Chapter 17

Filiz Yesilkoy and
Chad Ropp
Department of Electrical and Computer Engineering
University of Maryland
College Park
MD 20742, USA

Zach Cummins and Roland Probst
Fischell Department of Bioengineering
University of Maryland
College Park
MD 20742, USA

Edo Waks
Department of Electrical and Computer Engineering

and

Institute for Research in Electronics and Applied Physics (IREAP)
University of Maryland
College Park
MD 20742, USA

Benjamin Shapiro
Fischell Department of Bioengineering

and

The Institute for Systems Research (ISR)
University of Maryland
College Park
MD 20742, USA

Martin Peckerar*
Department of Electrical and Computer Engineering
University of Maryland
College Park
MD 20742, USA

E-mail: peckerar@umd.edu

Woodhead Publishing Series in Electronic and Optical Materials

1 **Circuit analysis**
J. E. Whitehouse
2 **Signal processing in electronic communications: For engineers and mathematicians**
M. J. Chapman, D. P. Goodall and N. C. Steele
3 **Pattern recognition and image processing**
D. Luo
4 **Digital filters and signal processing in electronic engineering: Theory, applications, architecture, code**
S. M. Bozic and R. J. Chance
5 **Cable engineering for local area networks**
B. J. Elliott
6 **Designing a structured cabling system to ISO 11801: Cross-referenced to European CENELEC and American Standards**
Second edition
B. J. Elliott
7 **Microscopy techniques for materials science**
A. Clarke and C. Eberhardt
8 **Materials for energy conversion devices**
Edited by C. C. Sorrell, J. Nowotny and S. Sugihara
9 **Digital image processing: Mathematical and computational methods**
Second edition
J. M. Blackledge
10 **Nanolithography and patterning techniques in microelectronics**
Edited by D. Bucknall
11 **Digital signal processing: Mathematical and computational methods, software development and applications**
Second edition
J. M. Blackledge

12 **Handbook of advanced dielectric, piezoelectric and ferroelectric materials: Synthesis, properties and applications**
Edited by Z.-G. Ye
13 **Materials for fuel cells**
Edited by M. Gasik
14 **Solid-state hydrogen storage: Materials and chemistry**
Edited by G. Walker
15 **Laser cooling of solids**
S. V. Petrushkin and V. V. Samartsev
16 **Polymer electrolytes: Fundamentals and applications**
Edited by C. A. C. Sequeira and D. A. F. Santos
17 **Advanced piezoelectric materials: Science and technology**
Edited by K. Uchino
18 **Optical switches: Materials and design**
Edited by S. J. Chua and B. Li
19 **Advanced adhesives in electronics: Materials, properties and applications**
Edited by M. O. Alam and C. Bailey
20 **Thin film growth: Physics, materials science and applications**
Edited by Z. Cao
21 **Electromigration in thin films and electronic devices: Materials and reliability**
Edited by C.-U. Kim
22 **In situ characterization of thin film growth**
Edited by G. Koster and G. Rijnders
23 **Silicon-germanium (SiGe) nanostructures: Production, properties and applications in electronics**
Edited by Y. Shiraki and N. Usami
24 **High-temperature superconductors**
Edited by X. G. Qiu
25 **Introduction to the physics of nanoelectronics**
S. G. Tan and M. B. A. Jalil
26 **Printed films: Materials science and applications in sensors, electronics and photonics**
Edited by M. Prudenziati and J. Hormadaly
27 **Laser growth and processing of photonic devices**
Edited by N. A. Vainos
28 **Quantum optics with semiconductor nanostructures**
Edited by F. Jahnke
29 **Ultrasonic transducers: Materials and design for sensors, actuators and medical applications**
Edited by K. Nakamura

30 **Waste electrical and electronic equipment (WEEE) handbook**
Edited by V. Goodship and A. Stevels
31 **Applications of ATILA FEM software to smart materials: Case studies in designing devices**
Edited by K. Uchino and J.-C. Debus
32 **MEMS for automotive and aerospace applications**
Edited by M. Kraft and N. M. White
33 **Semiconductor lasers: Fundamentals and applications**
Edited by A. Baranov and E. Tournie
34 **Handbook of terahertz technology for imaging, sensing and communications**
Edited by D. Saeedkia
35 **Handbook of solid-state lasers: Materials, systems and applications**
Edited by B. Denker and E. Shklovsky
36 **Organic light-emitting diodes (OLEDs): Materials, devices and applications**
Edited by A. Buckley
37 **Lasers for medical applications: Diagnostics, therapy and surgery**
Edited by H. Jelínková
38 **Semiconductor gas sensors**
Edited by R. Jaaniso and O. K. Tan
39 **Handbook of organic materials for optical and optoelectronic devices: Properties and applications**
Edited by O. Ostroverkhova
40 **Metallic films for electronic, optical and magnetic applications: Structure, processing and properties**
Edited by K. Barmak and K. Coffey
41 **Handbook of laser welding technologies**
Edited by S. Katayama
42 **Nanolithography: The art of fabricating nanoelectronic and nanophotonic devices and systems**
Edited by M. Feldman
43 **Laser spectroscopy for sensing: Fundamentals, techniques and applications**
Edited by M. Baudelet
44 **Chalcogenide glasses: Preparation, properties and applications**
Edited by J.-L. Adam and X. Zhang
45 **Handbook of MEMS for wireless and mobile applications**
Edited by D. Uttamchandani
46 **Subsea optics and imaging**
Edited by J. Watson and O. Zielinski

47 **Carbon nanotubes and graphene for photonic applications**
Edited by S. Yamashita, Y. Saito and J. H. Choi
48 **Optical biomimetics: Materials and applications**
Edited by M. Large
49 **Optical thin films and coatings**
Edited by A. Piegari and F. Flory
50 **Computer design of diffractive optics**
Edited by V. A. Soifer
51 **Smart sensors and MEMS: Intelligent devices and microsystems for industrial applications**
Edited by S. Nihtianov and A. L. Estepa
52 **Fundamentals of femtosecond optics**
S. A. Kozlov and V. V. Samartsev
53 **Nanostructured semiconductor oxides for the next generation of electronics and functional devices: Production, properties and applications**
S. Zhuiykov
54 **Nitride semiconductor light-emitting diodes (LEDs): Materials, technologies and applications**
Edited by J. J. Huang, H. C. Kuo and S. C. Shen
55 **Sensor technologies for civil infrastructures**
Volume 1: Sensing hardware and data collection for performance assessment
Edited by M. Wang, J. Lynch and H. Sohn
56 **Sensor technologies for civil infrastructures**
Volume 2: Applications in structural health monitoring
Edited by M. Wang, J. Lynch and H. Sohn
57 **Graphene: Properties, preparation, characterisation and devices**
Edited by V. Skákalová and A. B. Kaiser
58 **Handbook of silicon-on-insulator (SOI) technology**
Edited by O. Kononchuk and B.-Y. Nguyen
59 **Biological identification: DNA amplification and sequencing, optical sensing, lab-on-chip and portable systems**
Edited by P. Schaudies
60 **High performance silicon imaging: Fundamentals and applications of CMOS and CCD sensors**
Edited by D. Durini
61 **Nanosensors for chemical and biological applications: Sensing with nanotubes, nanowires and nanoparticles**
Edited by K. C. Honeychurch

62 **Composite magnetoelectrics: Materials, structures, and applications**
G. Srinivasan, S. Priya, and N. Sun
63 **Quantum information processing with diamond: Principles and applications**
Edited by S. Prawer and I. Aharonovich
64 **Advances in nonvolatile memory and storage technology**
Edited by Y. Nishi
65 **Laser surface engineering: Processes and applications**
Edited by J. Lawrence, C. Dowding, D. Waugh, J. Griffiths

Preface

No one ever thought that Moore's law would go on forever, that feature widths would continue to decrease by a factor of two every six years, or that the number of gates on a chip would double every six years or so. But these rates have never slowed down; if anything, they have become slightly faster. And now, with features *in production* just a few dozen atoms wide, the end is almost in sight. How will we deal with features ten atoms wide? What if they're not always ten, but sometimes nine and sometimes eleven? Will we call the two in the middle the thumbs, and the two at the edges the pinkies? Suppose gates shrink to just a few atoms. What will the wires connecting them look like?

The paths we have taken to get this far are fascinating, often using effects that everyone knew about, but were too constrained by long-standing inhibitions to consider practical. For example, immersion microscopy, in which a liquid between the objective and the work-piece is used to improve resolution, was probably older than anyone working in lithography. But for many years the worst crime a lab technician could commit was to allow anything to touch the surface of a resist-coated wafer. Yet now, exposing wafers under water is the mainstream method for fine line patterning.

As another example, contact printing was, and still is, a method for obtaining medium levels of resolution. Combined with the shorter wavelengths of X-rays, it promised resolutions below 100 nm. But the inhibition against touching mask to wafer in a production environment was so strong that even a 10 micron gap between the mask and wafer was considered dangerously small. This was one of the factors that ultimately led to the abandonment of X-ray lithography. Yet now, imprinting, in which the mask is literally pushed into the resist, is a major contender for the next generation of lithography.

As a final example, it has been long recognized that isolated features can be made narrower than densely packed ones. The trick is to control the developing, so as not to make the feature width zero, i.e., not to lose the feature completely. In addition, the use of double exposures with stencil masks was well known. But the tighter control and the longer exposure time did not seem suitable for a production environment. Nevertheless, the use of a first exposure for every other feature, and a second exposure for the remaining features, is also a major contender for the next generation of lithography.

This book is intended as a guide to the novice reading technical journals or facing the complexities of a conference dealing with lithography or nano-manufacturing. The novice may be a graduate student to whom everything is new, or he/she may be an experienced worker in a peripheral field. The authors, all experts in their fields, were instructed to give enough background information to enable novices to understand, and appreciate, new papers in those fields. My thanks to them for their patience in accepting this challenge.

The future belongs to the novice, armed with the knowledge of the past, but unhampered by its inhibitions. How will Moore's law end? I look forward to finding out.

M. Feldman
August, 2013

1

Optical projection lithography

B. W. SMITH, Rochester Institute of Technology, USA

DOI: 10.1533/9780857098757.1

Abstract: Optical projection lithography has been the predominant method of micro- and nano-patterning for most semiconductor and nanotechnology applications. This chapter addresses the approaches, systems, and materials that have been used, as optical lithography has enabled patterning from the micrometer scale down to the nanometer scale. The technology involved with the development of optical systems and imaging methods is discussed, including the optics, sources, photomasks, illumination, and techniques for resolution enhancement.

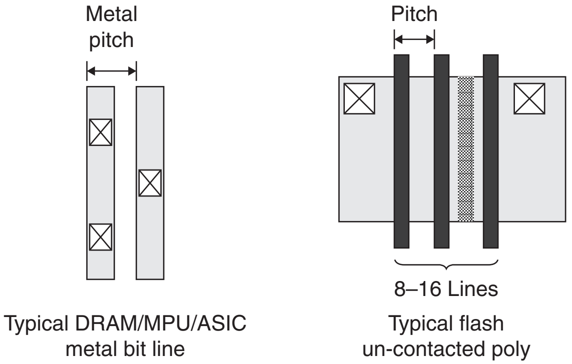
Key words: optical lithography, microlithography, projection lithography, DUV lithography.

1.1 Introduction

Optical projection lithography has been a key enabler as the functionality in nanoelectronic, nanophotonic, and nanobiological devices has increased. To understand the role that optical lithography plays in achieving higher resolution and greater density devices, the silicon semiconductor integrated circuit (IC) is most often looked to. This chapter provides a description of the underlying technology of optical lithography for nanoscale patterning as it has evolved through generations of the IC into the deep sub-wavelength, nanometer-level technology of today. A review of the fundamentals of projection UV optical lithography is given, followed by in-depth description of the technology used to advance IC device performance to near theoretical limits. Application of optical lithography to new generations of nanoelectronic devices, as well as into new nanoscale frontiers, will continue to make use of this core enabling technology.

1.2 Lithography technology and trends

Optical lithography has been the dominant method of patterning for semiconductor device manufacturing since the inception of the IC. Although other lithography approaches, such as electron beam, X-ray, ion-beam, imprint, and extreme ultra-violet (EUV) have been pursued (and are


Metal pitch
Pitch
8–16 Lines
Typical DRAM/MPU/ASIC metal bit line
Typical flash un-contacted poly

1.1 Typical critical dimension (CD) and pitch designations defined for semiconductor geometry in dynamic random access memory (DRAM), microprocessors (MPU), application specific integrated circuits (ASIC), and flash memory. (Reference 1.)

discussed in later chapters), the needs of high volume, high resolution chip production have historically been met mostly by optical methods. By looking at current and past editions of the International Technology Roadmap for Semiconductors (ITRS), variations to the basic optical lithography theme are evident.[1] Throughout the years, improvements have included the use of shorter wavelengths, larger numerical apertures (NA), customized illumination, phase-shifting masks, mask correction methods, immersion lithography, constrained geometry, and co-optimization of enhancement methods to achieve the resolution, focal depth, and overlay control required for nanoscale device patterning.

Technology trends that follow roadmaps like the ITRS often track the needs and progress of various IC types, such as memory devices, with high regularity and duty ratios near 1:1 (line-to-space ratio), and microprocessors with less regularity and more sparse duty ratios. Throughout the generations of these devices, dynamic random access memory (DRAM) and, more recently, flash memory have driven dense pattern geometry to the smallest levels. IC geometry size is often scaled in terms of the *half-pitch*. The concept of half-pitch is not based on a single geometric feature but instead derived to establish a comparative metric that could be used for devices' various geometry density. It is defined simply as one half of the sum of a line and a space. Although the less dense patterns for microprocessor devices are generally smaller than those for memory devices, their corresponding half-pitch has conformed to less aggressive scaling. Figure 1.1 shows how typical gate geometry defines critical dimension (CD) lines and pitch, where the CD is the smallest feature of interest in a particular device level.

A summary of IC trends and the corresponding lithography technology is shown in Table 1.1, where device 'generations' are listed together with

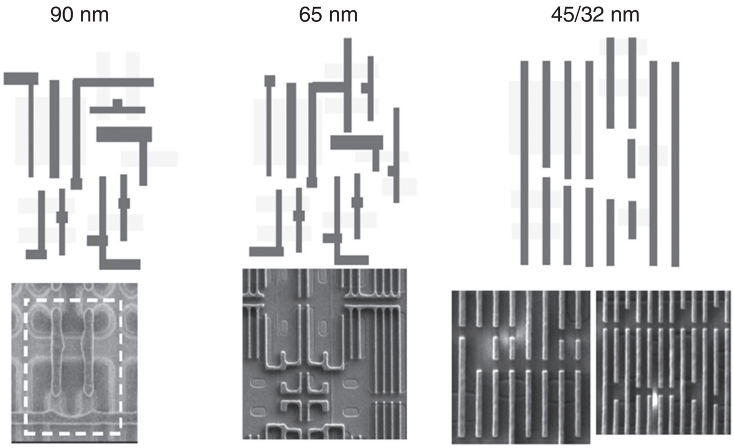
Table 1.1 Lithography methods used for 40 years of IC generations showing memory devices (DRAM and Flash) and microprocessor (MPU)

	DRAM/Flash[a] hp (nm)	MPU hp	Lithography Method
1981	2000	4000	g/i-line contact
1984	1500	3000	g/i-line steppers
1987	1000	2000	g/i-line
1990	750	1500	g/i-line
1993	500	1000	i-line
1995	350	750	i-line to 248 nm DUV
1997	250	350	248 nm DUV scanners
1999	180	250	248 nm DUV
2001	130	180	248 to 193 nm DUV
2003	90	145	193 nm DUV
2005	65	115	193 nm DUV to immersion
2007	45	90	193 nm immersion
2009	38	54	Immersion DP
2011	28	38	Immersion DP
2013	23	27	Immersion DP/MP
2015	18	25	Immersion MP to EUV
2017	14	20	Immersion MP to EUV
2019	11	16	EUV/ML2/Imprint/DSA/MP
2021	9	11	EUV/ML2/Imprint/DSA/MP

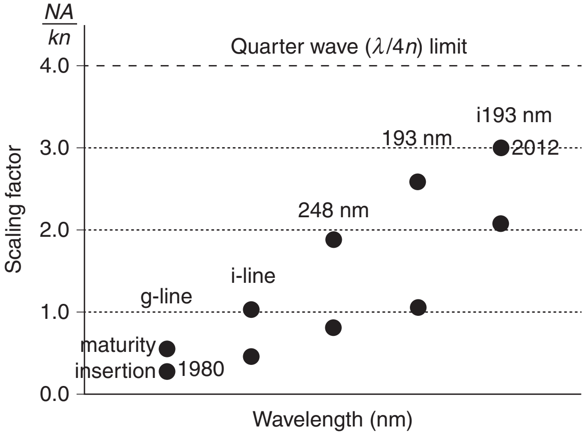
[a] DRAM represented minimum hp until 90 nm generation where Flash takes over.

corresponding half-pitch dimensions and lithographic exposure wavelengths. As generations shrink toward device features that are a fraction of the exposing wavelength, constraints have been placed on the structure, orientation, and variety of geometry involved so that a given wavelength technology can be used as long as possible. While earlier IC generations may have allowed for nearly unrestricted geometry choices, current and future devices demand tight constraints. This can be seen for example in Fig. 1.2, where the evolution of a microprocessor design is shown for generations from 90 to 32 nm.[2] As shown, 90 nm and larger generations had allowed design layout that only constrained gates to one direction with few other limitations. The evolution to 65 nm added design rules for a variety of allowable pitch values. For 45 and 32 nm design rules, all gates are constrained to one direction with a single pitch. This trending toward more regular, single pitch 'grating'-like designs has allowed for the extension toward higher density and smaller device geometry.

As Table 1.1 shows, the progress of lithography has resulted in the scaling of resolution with wavelength (or vice versa if thought of in terms of the demands of resolution). Early generation device geometry was many



1.2 Microprocessor design evolution and resulting wafer level patterning for generations from 90 to 32 nm. (*Source*: From Schenker R, Singh V and Borodovsky Y (2010), 'The role of strong phase shift masks in Intel's DFM infrastructure development,' SPIE Design for Manufacturability through Design-Process Integration IV, 7641.)



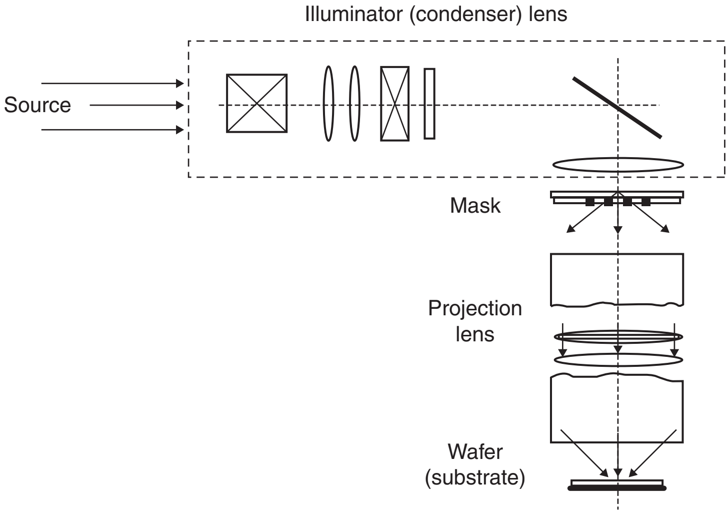
1.3 Wavelength scaling trends of optical lithography generations over the past several years where the scaling factor is the measure of resolution with respect to wavelength. (*Source*: From 'The saga of lambda: spectral influences throughout lithography generations,' B.W. Smith, Proc. SPIE Advances in Resist Materials and Processing Technology XXIX, 8325, 2012.)

times larger than the exposing wavelength, while current and future generations have structures at fractional wavelength sizes. Figure 1.3 is a plot showing the wavelength scaling trends over the past several decades.[3] The plot shows how half-pitch resolution has scaled with wavelength at insertion and maturity points for several generations. The scaling factor (NA/k_1n) is a measure of resolution with respect to wavelength, with a limit at a quarter-wavelength ($\lambda/4n$). Here, λ is wavelength, NA is the lens numerical aperture, k_1 is a process scaling factor, and n is the image media refractive index. Additional discussion of these terms is provided later in this chapter. Immersion lithography has achieved an impressive scaling factor of a third-wavelength ($\lambda/3n$). The timing of the technology advances necessary for such scaling has been tied to the incremental learning that has occurred during these wavelength transitions.

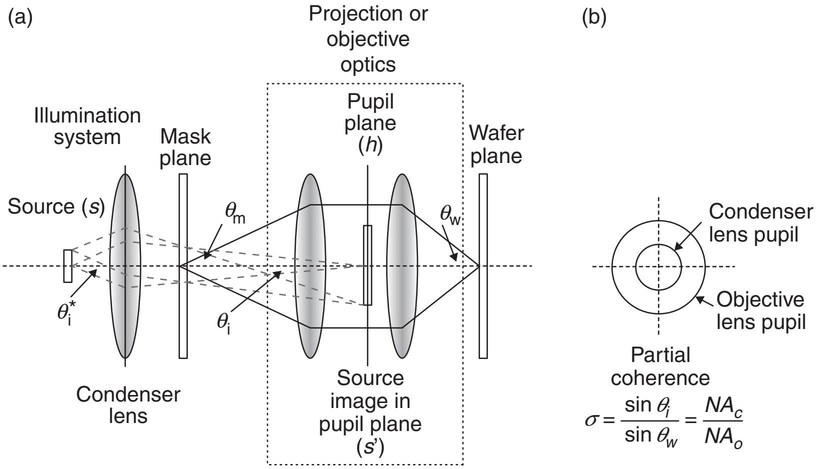
1.3 Fundamentals of optical lithography

The basic components of an optical projection lithography system are shown in Fig. 1.4. These include a source of uniform ultraviolet (UV) radiation, a photomask, a projection lens (also referred to as an objective lens), and an image plane. Projection lithography systems like this are most often configured using Köhler illumination, in the same manner as a conventional optical microscope. In a Köhler illuminated system, the source is imaged into the pupil plane of the objective lens and the object (the mask) is imaged at the image (the wafer) plane. This is shown in Fig. 1.5a, where rays are traced through the system to show these image locations. There are several NA locations in the configuration corresponding to object and image sides of the condenser lens and the projection lens. Specifically, θ_i and θ_i^* are the half-collection angles associated with the condenser lens, where the associated NAs in a media with refractive index n are NA$_i$ = $\sin\theta_i$ and NAi* = $\sin\theta_1^*$. The collection angles associated with the projection lens are θ_m (on the mask side of the lens) and θ_w (on the wafer side). The corresponding NAs are NA$_m$ = $n\sin\theta_m$ and NA$_w$ = $n\sin\theta_w$. In a conventional projection system, air is the image medium and NA$_w$ is simply $\sin\theta_w$. For an immersion lithography system, NA$_w$ is increased by the refractive index of the immersion fluid, which is about 1.44 at 193 nm with water. The reduction value of the projection lens is the ratio of NA$_w$/NA$_m$.

Through the manipulation of these apertures relative to each other, control of the spatial coherence of the system is possible. Spatial coherence in a projection system is defined as the relationship of the phase of light as it propagates from the illuminator toward a mask. If there is no angular distribution to the light (which can be achieved by using a very small aperture in the illuminator), there is no distribution of the phase of the light and it can be considered spatially coherent. As the illuminator NA is increased, the angular distribution of the light directed toward the mask is also increased, leading to a decrease in the phase coherency. This is referred to as partially coherent illumination. If the aperture is at a maximum, allowing all possible angles of illumination at the mask, the condition becomes one of


Illuminator (condenser) lens
Source
Mask
Projection lens
Wafer (substrate)

1.4 Basic components of a projection lithography system, including the source, mask, projection lens, and substrate.


(a)
Projection or objective optics
Illumination system
Mask plane
Pupil plane (h)
Wafer plane
Source (s)
θ_m
θ_w
θ_i^*
θ_i
Condenser lens
Source image in pupil plane (s')
(b)
Condenser lens pupil
Objective lens pupil
Partial coherence
$\sigma = \frac{\sin \theta_i}{\sin \theta_w} = \frac{NA_c}{NA_o}$

1.5 Optical elements of projection imaging system illuminated with Köhler illumination. Shown are (a) the object and image locations of the illumination and projection optics, and (b) the image of the condenser lens pupil in the objective lens.

spatial incoherence. In order to quantify these levels of spatial coherence, the illuminator NA needs to be measured against the ability of the projection system to collect these angles. In other words, a ratio of the NA_c of the illuminator to the NA_o of the projection lens will define the level of spatial coherence. This is known as the partial coherence factor (σ), which is depicted in Fig. 1.5b where:

$$\sigma = \frac{n \sin \theta_i}{n \sin \theta_w} = \frac{\mathrm{NA}_c}{\mathrm{NA}_o}$$

When the ratio of NA_c to NA_o is small, mask illumination is more coherent. As the ratio approaches unity, the illumination is more incoherent. This source property is important as it influences the distribution and collection of the light diffracted by a mask. In doing so, it will also have a strong influence on image quality and ultimate resolution, as can be demonstrated from some basic examples.

Consider first a mask comprising a simple binary grating with pitch (p) of lines and spaces illuminated at normal incidence using light of wavelength λ. Normal incidence results in spatially coherent illumination, which is achieved with a condenser lens NA_c that is effectively zero (like a point source). Light is diffracted by the mask features to produce diffracted orders distributed according to the grating formula:

$$\sin \theta_m = \pm \frac{m\lambda}{p},$$

where m is the integer order number. If all of these diffracted orders can be collected by an objective lens, a perfect image can be produced at the image plane. Realistically, however, there will be a limit to the number of orders that can be collected, which is most often small. Figure 1.6a depicts the situation for the collection of just the primary orders, the 0th and ±1st. The resulting image that will be projected toward the image plane will have lost all high frequency content and be imaged as a biased sinusoidal function, as shown in Fig. 1.6b. Such an image can be sufficient if it is recorded into a high contrast photoresist, which can create high aspect relief features if there is sufficient modulation in the optical image. The goal of optical lithography therefore becomes one of imaging smaller dimensions while collecting sufficient diffraction energy so that the response of a photoresist is usable. Although higher order diffraction energy is generally not needed, methods to distribute diffraction energy in ways that allow for its collection for finer pitch features are desirable.

A basic approach used to manipulate the distribution of diffracted light is through the control of the partial coherence factor, σ. Instead of using normally incident illumination from a small condenser lens NA_c, as shown in Fig. 1.6, higher values of partial coherence will spread diffracted energy over a range of angles, as shown in Fig. 1.7. This is achieved by using a larger NA_c to NA_o ratio. For example, if Fig. 1.5b is plotted on $\sin\theta$ axes, the σ ratio depicted is about 0.5. The illumination of a mask with this angular distribution of light would not result in discrete diffraction orders as shown in

Fig. 1.6a. Instead, each order would have an angular spread equivalent to that for the illumination, as shown in Fig. 1.7a. When the resulting diffraction energy is captured by an objective lens, a lens of lower NA_o can collect 1st diffraction order energy than would be necessary for coherent illumination. This is shown in Fig. 1.7b where three values of partial coherence show that a smaller objective lens is needed with increasing σ values. However, instead of choosing a smaller objective lens NA, it is desirable to image smaller features with a given NA. Higher partial coherence values are therefore often used for many applications where regular fine structure needs to be imaged.

1.4 Image evaluation

A primary goal of a lithography system is to achieve high resolution over a sufficient field and focal depth. Device field requirements have led to lens field diameters on the order of 2–3 cm. Focal depth, or depth of focus (DOF) must be large enough to produce usable imaging throughout an entire photoresist layer. DOF requirements are tied to the dimensional requirements of a particular device technology, and may be 100 nm or greater. The fundamental relationships between resolution, DOF, and optical system parameters can be represented in the Rayleigh equations:

$$\mathrm{hp}(\mathrm{min}) = \frac{k_1 \lambda}{\mathrm{NA}},$$

and

$$\mathrm{DOF} = \frac{\pm k_2 \lambda}{n \sin^2 \theta},$$

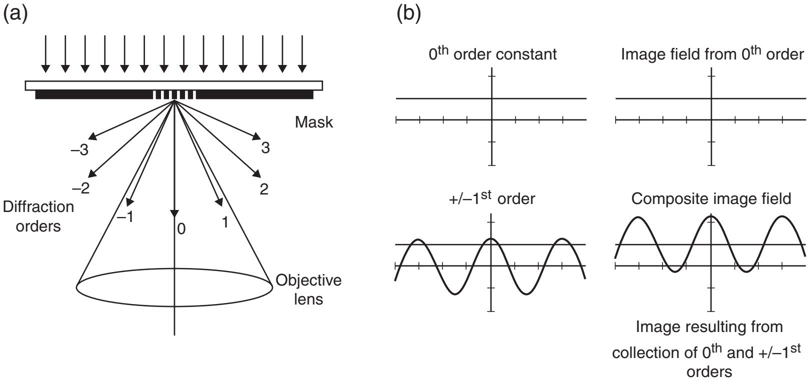
where k_1 and k_2 are process-dependent factors tied to the capability of the lithography system and photoresist process, and n is the refractive index.

1.4.1 Image resolution

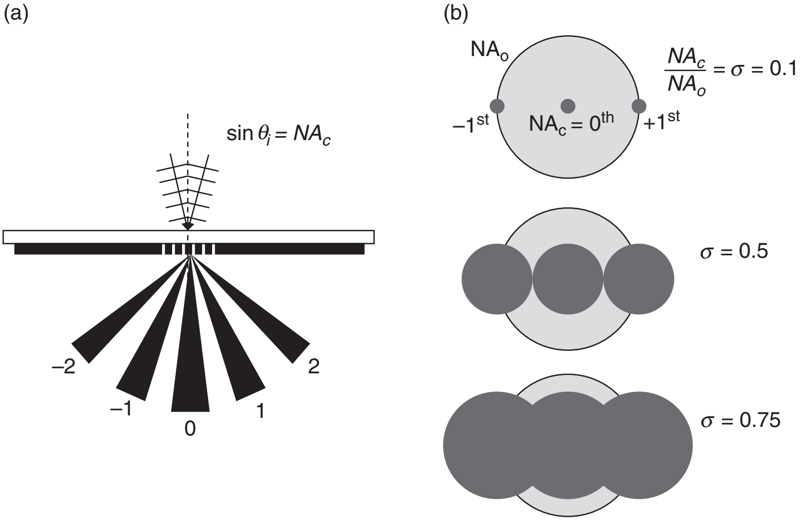
The Rayleigh resolution equation can be derived using Figs 1.6 and 1.7 together with the grating equation. For coherent illumination (Fig. 1.6), where diffraction orders are located at multiples of $\pm m\lambda/p$, the objective lens NA_o must be sufficiently large so that it collects first diffraction order energy, or:

$$\mathrm{NA}_o \geq \frac{\lambda}{p}.$$

Through rearrangement to solve for the minimum half-pitch, this becomes:


(a)
Mask
−3
3
−2
2
Diffraction orders
−1
0
1
Objective lens
(b)
0^{th} order constant
Image field from 0^{th} order
+/−1^{st} order
Composite image field
Image resulting from collection of 0^{th} and +/−1^{st} orders

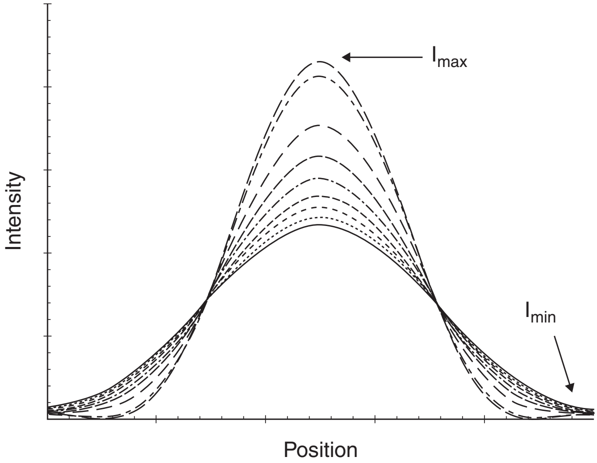
1.6 Projection imaging with coherent illumination where (a) single discrete diffraction orders are collected, and (b) the resulting image electric field is a biased sinusoid.


(a)
$\sin\theta_i = NA_c$
−2
2
−1
1
0
(b)
NA_o
$\frac{NA_c}{NA_o} = \sigma = 0.1$
−1^{st}
$NA_c = 0^{th}$
+1^{st}
$\sigma = 0.5$
$\sigma = 0.75$

1.7 Projection imaging with partially coherent illumination where (a) distributed diffraction orders are collected and (b) the result is the ability to collect more diffraction energy in the objective lens.

$$\mathrm{hp}(\min) = \frac{0.5\lambda}{\mathrm{NA}_o}.$$

For partial coherence (Fig. 1.7), the objective lens NA can be smaller than λ/p by the amount of diffraction order 'spread' imparted by the angular distribution of illumination, $\sin\theta_i$:



1.8 Aerial image of a space feature or mask opening, showing minimum and maximum intensity values through defocus.

$$\mathrm{NA}_o \geq \frac{\lambda}{p} - \sin\theta_i$$

By dividing both sides of the equation by NA_o and substituting NA_c for $\sin\theta_i$, and σ for $\sin\theta_i/\mathrm{NA}_o$, the minimum half-pitch can be solved for as:

$$\mathrm{hp}(\min) = \frac{0.5\lambda}{(\sigma+1)\mathrm{NA}_o},$$

which is a more general solution for any condition of partial coherence.

1.4.2 Image quality

As the partial coherence factor is increased, the minimum resolvable half-pitch can decrease. This does not necessarily correlate to image quality, which will decrease as the fractional amount of diffraction order energy decreases from the spread of diffraction orders. Image quality metrics are thus required to determine how usable an image is. Practical resolution is related to image quality as much as it is related to the ultimate limits of the Rayleigh equation. Consider the image of a space feature (i.e. a clear opening or a trench), as shown in Fig. 1.8. The image exists as an intensity distribution along a plane $I(x)$, known as the aerial image. The plot shows how image degradation, such as defocus, will degrade an aerial image and reduce its quality. Several metrics of image quality could be used to quantify this degradation. Image modulation (also known and image contrast) measures the normalized ratio of minimum to maximum intensity as:

$$\text{Image modulation}\,(M) = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}}$$

For a given system, the change in modulation relative to density of the geometry that is imaged (or its spatial frequency) is known as the modulation transfer function (MTF). Modulation and MTF are common image metrics used in many optical situations. Strictly speaking, modulation is only linear with incoherent illumination (where image intensity will add). It is also a difficult metric to use for mask features with limited periodicity where there are no well-defined maximum and minimum intensity values. Inspection of the aerial image plot of Fig. 1.8 shows the opportunity to utilize a quality metric that measures the degradation of the slope of the aerial image, $\Delta I(x)/\Delta x$. Such a slope measurement should, however, take into account the logarithmic response of the detector (which is photoresist in this case). Taking the log of an aerial image prior to measuring the sidewall slope leads to a response metric that can behave linearly with varying input. Specifically, the image log slope (ILS) is a measure of the slope of the log of the aerial image:

$$\text{ILS} = \left.\frac{\mathrm{d}\ln I(x)}{\mathrm{d}x}\right|_{\text{Mask Edge}}$$

which is measured at the position of the transitional edge of the targeted feature (the *Mask Edge*). The units of ILS are reciprocal distance (such as nm^{-1}). The product of ILS and feature size provides a unitless metric, which provides an indication of image quality regardless of feature size, known as the normalized image log slope (NILS):

$$\text{NILS} = \text{ILS} \times \text{CD} = \frac{\mathrm{d}\ln I(x)\,\text{CD}}{\mathrm{d}x}$$

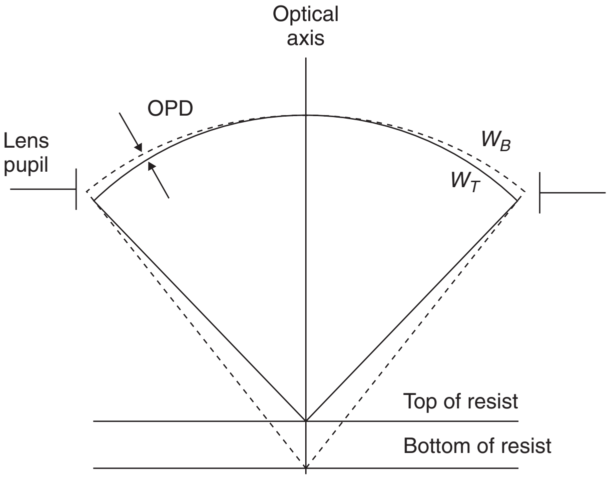
1.4.3 Depth of focus

As seen in the Rayleigh DOF equation, focal depth falls off linearly with shorter wavelength and quadratically with the sine of the collection angle.

For air as an image medium (with a refractive index of 1.0), the denominator of the DOF equation becomes NA^2, which shows how the optical properties that lead to higher resolution (wavelength and NA) are detrimental to focal depth. The process factor k_2 in the equation may be a value near 0.5 for a conventional high contrast photoresist, and values greater than that imply a resist that is more capable of utilizing a poorly modulated aerial image. The full usable depth of focus (UDOF) is often extracted from photoresist exposure tests through a range of focus and exposure conditions (a focus–exposure matrix). The UDOF is the full range of focus that leads to a feature of a CD that remains within a required specified size range (the CD specification) over an established range of exposure variation, the exposure latitude (EL). Common values for the CD specification and EL are ± 10% and ± 5% respectively. Using the UDOF derived from such exposure testing can lead to a solution for a process-specific k_2 value by solving the equation:

$$k_2 = \frac{\mathrm{UDOF} \times \mathrm{NA}^2}{2\lambda}$$

The relationship between DOF, NA, and wavelength can be visualized by comparing a wavefront propagating through an objective lens pupil focused at different positions at an image plane, as shown in Fig. 1.9. For two image plane locations (i.e. the top and the bottom of a photoresist layer), there exists two unique spherical waves in the lens pupil that would converge to focal points at each respective plane. For a single point at the object (mask) plane, centered on the optical axis, these are the two spherical waves that would be associated with focusing an image of that point at the top and the bottom of the resist. In other words, when the point is focused at the top of the resist, the wave W_T would need to exist in the lens pupil. When the point is focused at the bottom of the resist, the wave W_B would be necessary. The challenge for lithography is to achieve good focus throughout a depth sufficient to account for the photoresist thickness and any substrate topography. This is the DOF budget. The error associated with choosing a single focal plane and evaluating the image at other locations can be assessed by considering the optical path difference (OPD) between the corresponding spherical waves. This OPD increases with a larger pupil (and NA_o). Since the OPD is a result of the difference between two spherical waves, it will increase quadratically with pupil size. Hence the presence of NA^2 in the denominator of the Rayleigh DOF equation. Since this OPD is the distance traveled by light propagating through the lens pupil, wavelength also influences the impact of the OPD. Longer wavelength light experiences less wavelength error through a given physical path length than shorter wavelength light does. This explains why



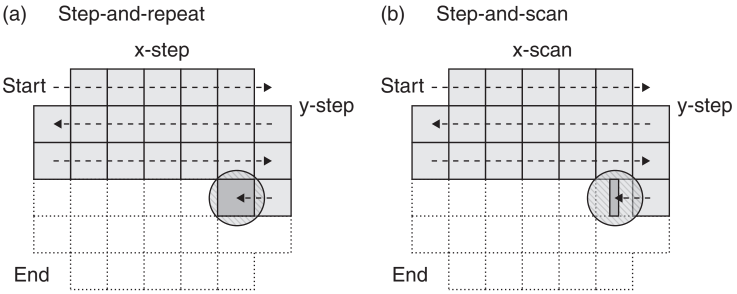
1.9 Comparison of the wavefronts corresponding to two focus positions in the substrate plane, where the OPD is the separation between them.

wavelength is in the numerator of the Raleigh DOF equation. Evaluating an optical imaging system from the perspective of the lens pupil is useful for understanding other phenomena as well, including aberrations that will be discussed in a subsequent section.

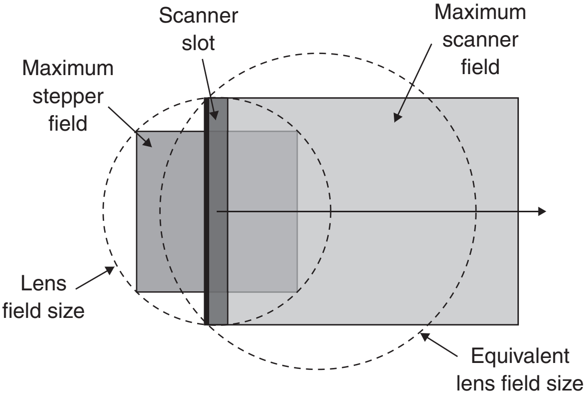
1.5 Projection lithography systems

Historically, optical lithography systems have been designed to image a photomask onto a substrate in various ways including contact, proximity, scanning, and stepping modes or variants thereof. Until the late 1990s, many projection lithography systems had been configured to operate in what is referred to as a step-and-repeat imaging mode. Commonly referred to as 'steppers,' these systems expose an arrayed matrix of IC chip fields by exposing one field at a time and stepping through rows and columns to cover an entire wafer substrate, as shown in Fig. 1.10a. More recently, a step-and-scan imaging mode (a step-and-scanner) has been employed where a row of chip fields is scanned in one direction, followed by a stage stepping motion to allow for the imaging of additional rows, as shown in Fig. 1.10b.

A main advantage of step-and-scanning over step-and-repeating is realized when considering the utilization of the projection lens field. This is shown in Fig. 1.11, where the lens field is depicted together with its utilization for a square stepper field and a slot scanner field. Since the scanner requires only a narrow portion of the lens field (about 5 mm × 26 mm for


(a) Step-and-repeat
x-step
Start
y-step
End
(b) Step-and-scan
x-scan
Start
y-step
End

1.10 Exposure sequencing for (a) a step-and-repeat lithography system, and (b) a step-and-scan lithography system.

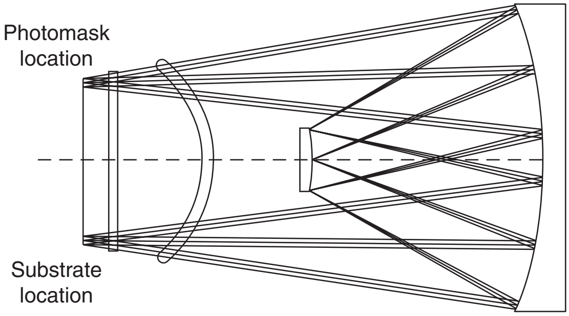

Scanner slot
Maximum scanner field
Maximum stepper field
Lens field size
Equivalent lens field size

1.11 Comparison of the maximum field sizes for stepper and scanner lithography systems using the same lens field size.

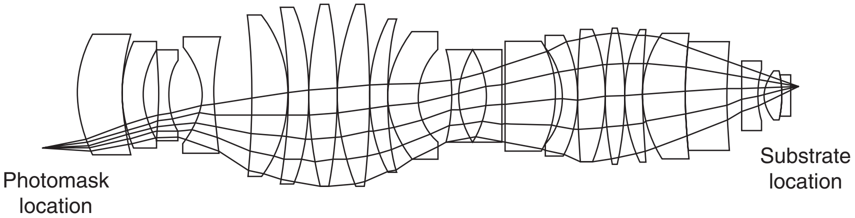
commercial systems), the imaging of larger chips compared to the stepper is enabled. This is an important advantage, as optical lithography systems have been driven to require both very large fields and very high NAs. Achieving both of these attributes is an optical challenge, and the field configuration for a slot scanner allows for a larger image field in a given projection lens without significant modifications to an optical design.

1.5.1 Projection lithography optics

The lenses used for optical lithography systems have comprised refractive elements (known as catoptric designs), reflective elements (known as dioptric designs), or combinations of the two (catadioptric designs). An early example of a mostly reflective lens design is that of the Perkin Elmer Micralign full wafer scanner. The objective lens for such a system is shown



1.12 Mostly reflective (catadioptric) optical design of a full wafer scanner.



1.13 Refractive (dioptric) design example of a reduction projection lithography lens.

in Fig. 1.12, where photomask and substrate each utilize one half of the full optical system diameter. The mask is illuminated with a long narrow arc slit of illumination, and the mask and substrate are scanned in unison across this slit. A 1:1 image of the mask is directed toward the wafer, which is complete when each is scanned across its entirety. This approach was a dominant imaging mode of the semiconductor industry in the 1980s for 3″ to 6″ wafers with resolution to about 1.5 μm.

Larger substrates and higher resolution were attainable using higher NA projection lenses and reduction projection imaging. Large field dioptric projection lenses with NA values up to 0.93 have been built for use in step-and-repeat and step-and-scan systems. An example of such a lens is shown in Fig. 1.13, where over 20 lens elements are combined in an all-refractive reduction projection lens. Reduction ratios for such lenses have been as large at 10× and as small as 2×. Problems associated with large field scaling and the implications for very large masks made 10× reduction impractical for high volume manufacturing. A small reduction ratio of 2× causes challenges for masks since feature sizes, mask defects, and dimensional control are nearly as difficult as for 1× reduction. The dominant reduction value for most projection lithography systems has converged to 4× where it remains today.

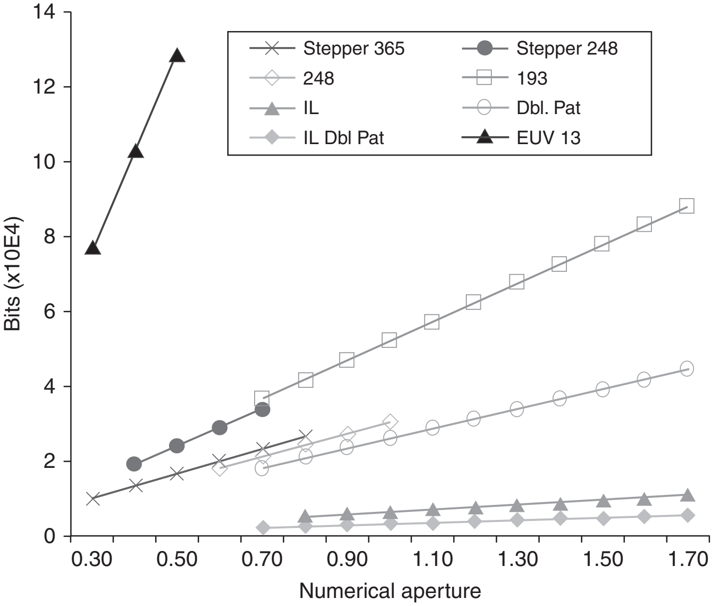
1.5.2 Numerical aperture and field size trends

The progress of IC technology has driven devices toward larger chip fields and smaller circuit features. This has placed increasing demands on the amount of information that must be transferred though an optical lithography system. A measure of the *information content* can be obtained by multiplying the image field of the system (in mm for example) and the spatial frequency of the image content (in cycles per mm for example, where 100 nm lines and spaces correspond to a spatial frequency of 500 cycles per mm). The product of these two system attributes gives a unitless indication of the information content of the lithography process. In optics, such a measure is referred to as an optical invariant or the *Lagrange invariant* (I), specifically:

$$I = \frac{\text{field radius} \times \text{NA}}{4k_1\lambda}$$

where k_1 is the Rayleigh resolution factor for the smallest feature (CD) of the device. This is a measure of the *bits* of information that the lens processes to image into the photoresist. As an example, a 365 nm 0.60 NA projection lens with a 35 mm image field diameter imaging 300 nm lines and spaces has a corresponding k_1 of 0.49 and can transfer 1.47×10^4 bits of information. The utilization of this field (as a square stepper field or a slot scanner field for example) will alter this accordingly. A leading edge 1.35 NA immersion lithography 193 nm scan-and-repeat system may print a 50 nm half-pitch CD with a 26 mm slot scanning 5.2×10^5 bits. With a throughput of 175 wafers/h, this corresponds to 1.13×10^{12} bits/s. Figure 1.14 is a comparative plot showing the information content of various lithography systems, including i-line (365 nm), 248, 193 nm, EUV (13.5 nm), interferometric lithography (IL), and double patterning. Details regarding EUV lithography can be found in Chapter 2. Double patterning and immersion lithography (necessary for NA values greater than 1.0) are addressed later in this chapter. Each line on the plot represents the approximate information content that could be produced using these technologies with systems that either have been built or have been contemplated.

As both the NA and field size requirements of lithography lenses have increased, scanning systems have been favored and design complexity has increased. For example, a high (0.80) NA 193 nm lithography system may

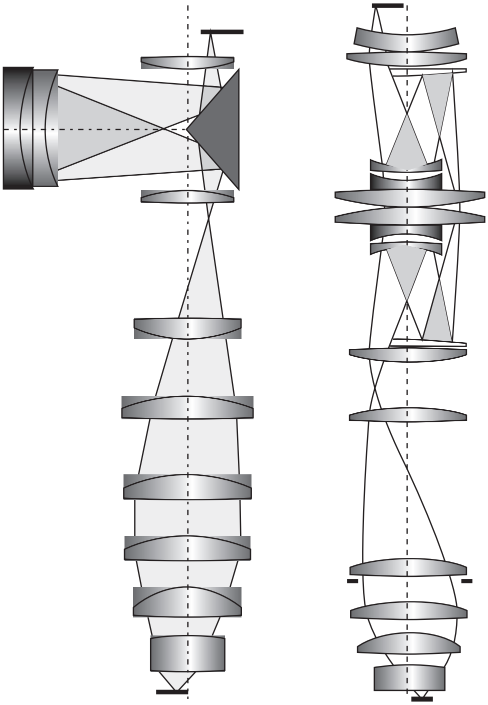


1.14 Comparative plot of the information content of various lithography systems measured as the Lagrange invariant in bits of image information.

have 25 or more refractive elements in a dioptric design. As acceptance angles (sinθ) are pushed above 0.90, the addition of reflective components is often necessitated for the reduction of aberrations across a large field, leading to the increased use of catadioptric lens designs. Figure 1.15 is an example of hyper-NA (NA > 1.0) 193 nm immersion lithography projection lens designs utilizing both refractive and reflective elements.[4]

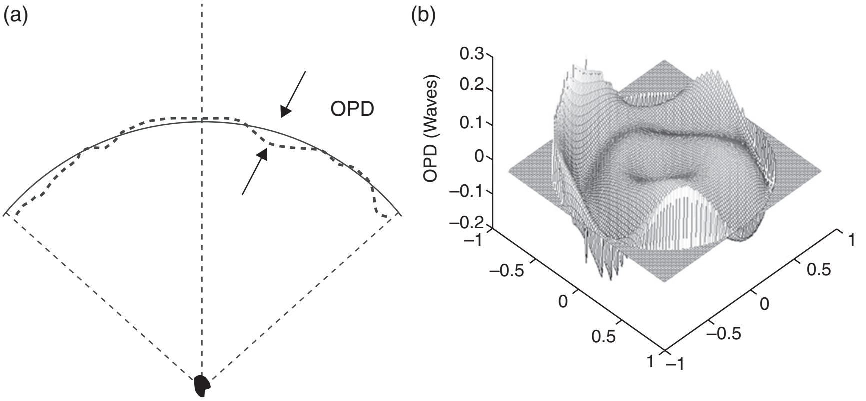
1.5.3 Aberrations, flare, and their influence

Image quality metrics such as modulation, MTF, ILS, and NILS are used to determine the fidelity of the aerial image resulting from the illumination of mask features and the collection of diffracted energy. Assuming no errors in the collection process, a resulting aerial image can be calculated based on diffraction theory alone. In reality, a lens system will introduce error to the collection and image formation process. This is known as lens aberration. Lens aberration can be evaluated and quantified from the perspective of the errors experienced as rays are traced through the optical system. Alternatively, the error can be evaluated in terms of the differences between an ideal wavefront in the lens pupil and that which can be achieved with a lens of a particular design and manufacturing process.



1.15 Examples of hyper-NA (NA > 1.0) optical projection systems for immersion lithography using catadioptric design approaches. (*Source*: From 'Immersion Lithography in Mass Production: Latest Results for Nikon Immersion Exposure Tools,' A. Hazelton, S. Wakamoto, K. Shiraishi, T. Fujiwara, H. Nagasaka, K. Nakano, S. Owa and T. Matsuyama, Sematech Immersion Symposium, 2007, http://www.sematech.org/meetings/archives/litho/8065/pres/TO-02%20Hazelton_Nikon%20final%20dist.pdf.)

This approach to quantifying the wavefront error as a deviation from an ideal was introduced in discussions about DOF and defocus through the concept of OPD. In this more general case of lens aberration, the OPD is the difference between an ideal spherical wavefront and that which can be achieved. Figure 1.16 depicts this, where the OPD varies across a lens pupil and will result in the blurring of a point object in a way that is characteristic of the particular deformation present in the lens. The OPD is also plotted for a lens pupil, showing various forms of symmetry and asymmetry as deviation from an ideal flat response that would result from a perfect lens. Characteristics of this deformation can be split into categories of aberration type, such as symmetrical defocus and spherical aberration or asymmetrical coma, astigmatism, and three-foil aberration. Various mathematical



1.16 Wavefront variation in an objection lens pupil in the presence of aberration (a) in one dimension showing the OPD, and (b) over the entire lens pupil.

decompositions and normalizations exist, including the Zernike polynomial description. The Zernike normalization balances aberration types in terms of a set of coefficients that are orthogonal over the lens pupil using the general polynomial form:

$$W(\rho,\theta)=\sum_{i}^{\infty}\alpha_i Z_i(\rho,\theta)$$

where n and m are positive integers ($n - m \geq 0$ and even), c_{nm} and s_{nm} are aberration coefficients, and the radial polynomial R of degree n in terms of the normalized radial coordinate in the pupil plane (ρ), and where i is a polynomial ordering index, ρ and θ are the normalized radial and angular coordinates of the lens pupil, $Z_i(\rho,\theta)$ is the ith Zernike polynomial function, and α_i is the Zernike polynomial coefficient. The Zernike radial polynomial function of degree n takes the form:

$$R_n^m(\rho)e^{\pm im\theta}$$

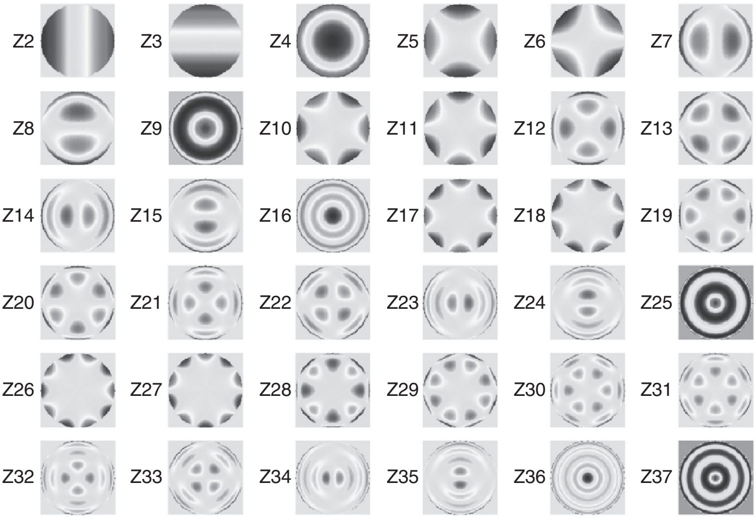
where n and m are positive integers ($n - m \geq 0$ and even).

The first nine Zernike polynomial terms are given in Table 1.2. These include tilt, defocus, astigmatism, coma, and spherical. The first term (Z_1 or piston) is a constant and represents an offset of the wavefront pupil wavefront.

Figure 1.17 is a plot of the first 37 Zernike polynomials. Various forms of symmetry allow for the identification of spherical terms (Z9, Z16, Z25, Z36,

Table 1.2 Zernike aberration terms for primary aberrations, where *r* is the pupil radius

Aberration	Phase aberration Φ (r, θ)
Spherical	Z_9 $[6r^4-6r^2+1]$
Coma	$Z_{7,8}$ $[3r^3-2r]\cos\theta$ $(\sin\theta)$
Astigmatism	$Z_{5,6}[r^2(\cos 2\theta)]$ $(\sin 2\theta)$
Defocus	Z_4 $[r^2-1]$
Tilt	$Z_{2,3}$ $[r\cos\theta]$ $(\sin\theta)$



1.17 Zernike polynomial expansion representations of individual aberration types including spherical terms (Z9, Z16, Z25, Z36, and Z37), astigmatism (Z5, Z6, Z12, Z13, Z21, Z22, Z32, and Z33), coma (Z7, Z8, Z14, Z15, Z23, Z24, Z34, and Z35).

and Z37), astigmatism (Z5, Z6, Z12, Z13, Z21, Z22, Z32, and Z33), coma (Z7, Z8, Z14, Z15, Z23, Z24, 34, and Z35), and others.

The influence of each of the aberration types will uniquely impact the imaging of various mask feature types. For example, spherical aberration, which is a radially dependent defocus effect, will cause features of different pitch to possess different best focus values. Asymmetry in focus-exposure-related behavior will also result. Astigmatism will produce different through-focus behavior for features oriented along one direction compared to another (for example X and Y orientation). Coma is an across-pupil aberration and will cause differences in the imaging of features on one side of a grouping compared to the other side. Image placement error (IPE) can result from coma aberrations as well.

The tolerable level of aberration in a lens will be a function of its application. For most optical applications, what is often referred to as the *quarter wave limit* (QWL) $\lambda/4$ is sufficiently good for a high quality lens. The QWL for a lens expresses the maximum OPD across the lens pupil, measured either in terms of variance (or RMS) or full range (or peak-to-valley). In a high quality lithography lens, the tolerable level of aberration is more likely below $\lambda/200$ (RMS). For a 193 nm lithography system, this corresponds to an astonishing 1 nm deviation between the ideal wavefront in the lens and the actual wavefront that can be achieved. Furthermore, the distribution for any single aberration term and the variation, as well as the variation of any one aberration across the field, is at the sub-nm level. Needless to say, the design, fabrication, and metrology processes involved in the manufacture of such optical systems present a great set of challenges.

Flare or stray light is a component of all optical systems, and will also lead to detrimental effects in image quality. Though flare can be measured in terms of its impact near the site of formation (short range) as well as the effect it has over a large area (far range), the general effect can quite simply be described in terms of its influence over general image quality. If a photoresist is exposed with an ideal image with an initial modulation M_i in the presence of flare (F), the image contrast will be reduced. A simple calculation of image contrast with flare (M_F) can be carried out as:

$$M_F = M_i \times (1 - F)$$

where the flare level is normalized to the initial source intensity. Flare in an optical lithography system is a result of reflections from interfaces in the projection lens (where 40+ layer surfaces may exist), surface scattering at each boundary, as well as reflections and scattering from the photomask, pellicle, and substrate. Anti-reflection coatings on optical surfaces reduce the specular reflection of optical surfaces and are an integral part of the lens design and manufacture process. Surface roughness leading to scatter is reduced as much as possible but becomes limited by the structural properties of the optical glass (as well as the ability to measure and reduce roughness as it approaches the nanometer level). Although earlier lithography systems with flare levels of several percent may have been tolerable for previous generation devices, leading edge ICs require that lithography systems operate with flare below 1%.

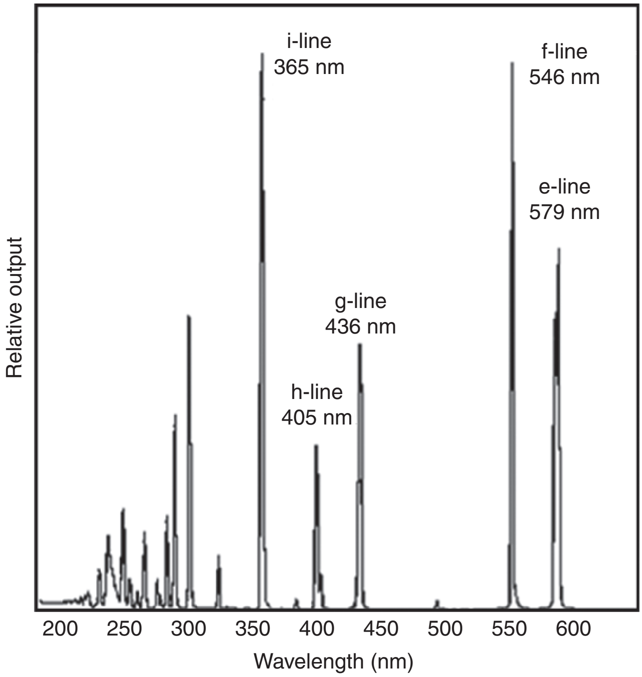
1.5.4 Alignment and overlay

Most micro-electrical, mechanical, or optical devices require the deposition and subsequent patterning of several film layers over a substrate in order to perform an intended function. In the case of a microelectronic IC, this layering comprises steps of thin film deposition, lithography, etching, thin film removal, and other steps that are replicated many times over. Modern nanoscale IC requires hundreds of separate process steps and tens of lithography masking levels over fields of several centimeters and wafer substrates 300 mm across. The accuracy of the alignment of one level to the next and the measurement of the resulting positional overlay (or registration) accuracy must be within a fraction of the minimum feature size. Errors in overlay will result in electrical problems such as variances in resistance, current leakage, and shorting, leading to yield issues. Improvements in overlay accuracy are achieved through optical, mechanical, design methods, detection schemes, and analysis means. A general specification for overlay control has been 25% of the device CD (3 sigma). For example, a 45 nm hp DRAM device process would require overlay control of 11 nm or better. Similarly, a 28 nm hp device would require 7 nm or better control. As device geometry progresses to where overlay control approaches a few nm, the ability to achieve such control may be the limiting factor to progress beyond a particular device generation.

1.6 Wavelengths for optical lithography

Diffraction scaling has driven optical lithography from portions of the electromagnetic spectrum known as the blue-UV (350–450 nm) to the mid-UV (or MUV, 300–350 nm), into the deep-UV (or DUV, 180–300 nm) and toward the vacuum-UV (or VUV, 100–180 nm). Mercury or mercury–xenon gas discharge lamps have traditionally been used to produce high power UV radiation in the blue and UV spectral regions. Emission lines are the result of transitions between electronic states, which are strong at 579 (e-line), 546 (f-line), 436 (g-line), 405 (h-line), and 365 nm (i-line). This is seen in the emission plot in Fig. 1.18. There are several weaker emission lines including a transition resulting in emission at 254 nm, which has also been used for lithography systems operating in the DUV.

Other sources have been considered or used for lithography at wavelengths down to 100 nm. Lyman-alpha (at 121 nm) and deuterium (at 160 nm) sources can emit at short-UV wavelengths but their power is generally too low to be practical. Starting with the argon (Ar_2) excited dimer (excimer) source at 126 nm, several excimer type sources were developed in the late 1970s including Kr_2 (148 nm), F_2 (157m,), the ArF rare-gas halogen excited complex source (an exiplex, but commonly referred to as an excimer) (193 nm), KrCl (222 nm), KrF (248 nm), XeCl (308 nm), and XeF (352 nm). The efficiency of excimer and exiplex sources allows them to be operated as lamps but higher power can be achieved through a narrow bandwidth as lasers. These lasers have come to be dominant for lithography exposure wavelengths in the DUV.

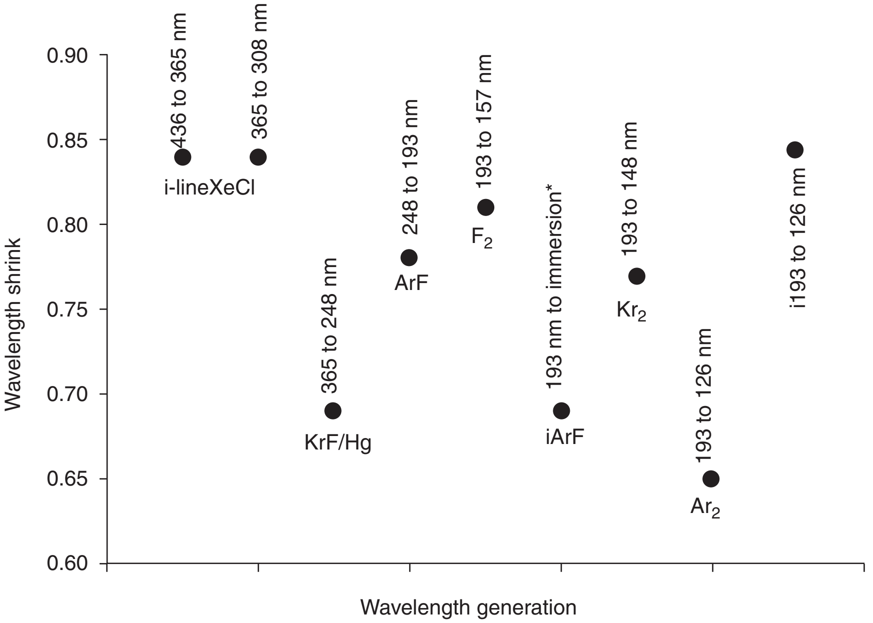


1.18 Spectral distribution of emission lines of a mercury vapor lamp, showing dominant lines including those used for lithography at 436 and 365 nm.

The transition to shorter wavelength must be justified with regards to the corresponding resolution potential. Simple calculations of wavelength scaling can be made, such as those shown in Fig. 1.19.[3] Here, the wavelength shrink corresponds to the ratio between old and new generations (and resulting in values less than one). It becomes apparent that there has been more potential in some transitions than others, where the step from i-line to 248 nm corresponds to a shrink better than 0.70. Such assessment shows how some lithography transitions were always difficult to justify, such as that from 193 nm (dry) to 157 nm lithography or any optical wavelength below 193 nm.

1.7 Lithography in the deep ultraviolet (UV)

The resolution improvements that are possible with shorter wavelength have driven lithography toward shorter DUV wavelengths. In moving lithography to wavelengths below 300 nm several material, design, and source factors have required consideration. Optics and optical materials issues have played a large role. The refractive influence of optical materials with wavelength,



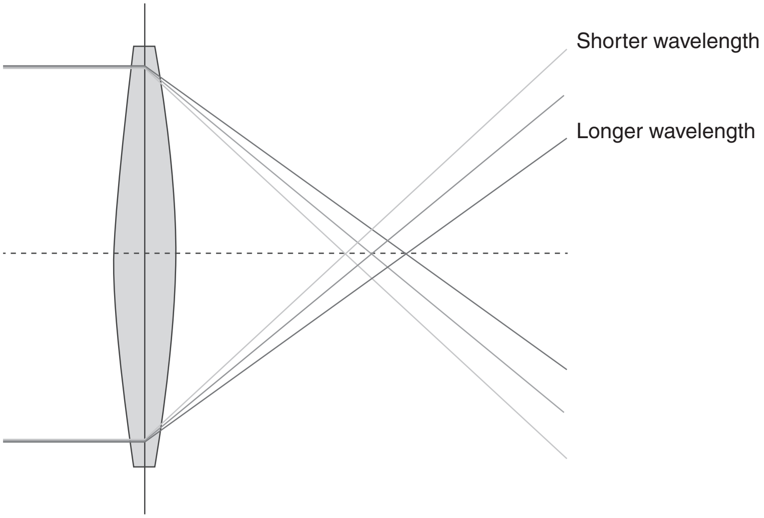
1.19 Fractional wavelength shrink from one lithography generation to the next, showing the potential improvement in resolution from wavelength reduction. (*Source*: From 'The saga of lambda: spectral influences throughout lithography generations,' B.W. Smith, Proc. SPIE Advances in Resist Materials and Processing Technology XXIX, 8325, 2012.)

defined through optical dispersion and manifesting as chromatic aberration, could be considered as one of the most important factors. The dispersion of an optical material is defined as the change in refractive index, Δn, for a change in wavelength, $\Delta\lambda$:

$$D=\frac{\Delta n}{\Delta \lambda}$$

For a refractive lens, chromatic aberration or color error will result in a variation in focus for different wavelengths. This is shown in Fig. 1.20 for a single positive lens element. The variation in focus between to wavelengths (or over a range of wavelengths) can be determined as:

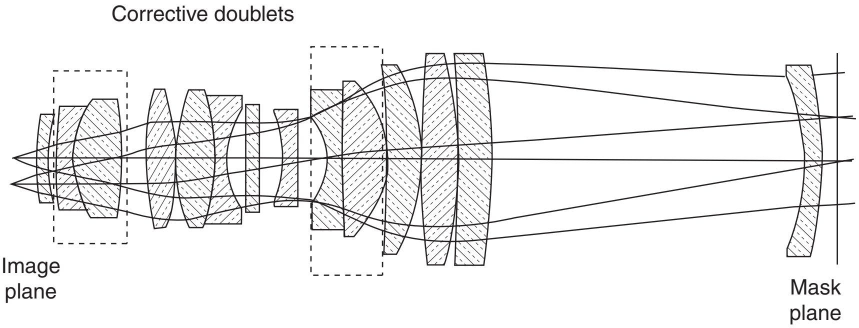
$$\delta f=\frac{f_0\Delta n}{n_0-1}$$



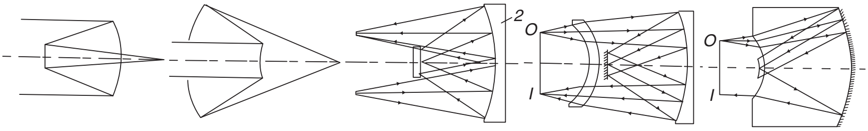
1.20 Chromatic aberration as the change in focus for different wavelengths for a refractive lens. In general, shorter wavelengths will focus closer to the surface of a positive lens as refractive index is increased.

where n_0 is the initial refractive index, and f_0 is the initial focal length. For a complex lens design, each element will be influenced by dispersion and, unless corrective measures are made, will only be well corrected for one wavelength or color. In a lithography projection lens, the demands on resolution, DOF, and image quality lead to spectral bandwidth requirements on the order of a picometers (pm) or better, measured as full width half maximum (FWHM) of the spectral distribution of the source. As seen from the spectrum for the mercury–xenon lamp of Fig. 1.18, the bandwidth at DUV wavelengths is more likely on the nm level or more, 1000× that which is necessary.

The use of cemented doublets in an optical design allows for the reduction of chromatic aberration through the combination of glass types with different refractive indices and dispersion values. This is seen in Fig. 1.21 for an i-line reduction projection lens that contains two doublets of different glass type within the design. At wavelengths below the absorption edge of traditional optical glass, few materials exist that can be combined for such correction. With fused silica as the main glass material for deep-UV optics, other means must be employed. Methods to reduce or eliminate the adverse effects of chromatic aberration include the use of a spectrally pure source (with bandwidth $\Delta\lambda$ in the picometer range in the case of lithography optics) or the use of entirely reflective optics (catoptric lenses) or mostly reflective optics (catadioptric lenses). Regarding the



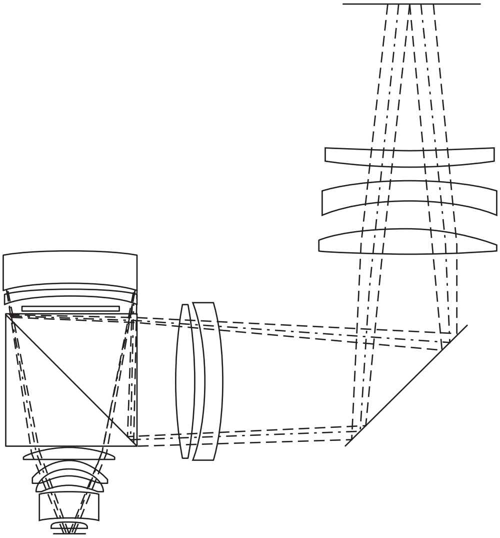
1.21 Use of cemented doublets in a refractive lens allow for the correction of chromatic aberration. This requires the use of different glass types in the lens design, which becomes difficult at DUV wavelengths.



1.22 Reflective (catoptric) design approaches leading to partially reflective (catadioptric) alternatives.

narrow bandwidth solution, this is problematic for excimer lasers where their high efficiency for populating metastable excited states leads to lower temporal coherence than most other lasers. At a natural bandwidth of 1 nm, additional correction methods are necessary. While spectral line-narrowing of excimer lasers has been explored since the late 1970s, using gratings, prisms, and/or etalons, the maturity of such techniques was not proven robust enough for high volume manufacturing until sometime later. Alternatively, the use of reflective optics, such as those shown in Fig. 1.22, exhibits no refractive chromatic aberration. Designs have evolved from the simple Schwarzschild objective to those used in commercial lithography systems. However, the use of reflective elements in a reduction projection lens (for stepper or step-and-scan use) did not reach high volume application until catadioptric designs, such as shown in Fig. 1.23, were introduced.[5]

Deep-UV lithography in the 248–54 nm range was primarily carried out using a catadioptric design like that in Fig. 1.23 and the 245 nm line of Hg, until sufficient success of the spectrally line narrowed excimer KrF laser. As shorter wavelengths were pursued, the 193 nm output from the ArF laser proved dominant over others. Issues involving the bandwidth requirement of high NA catadioptric lenses allowed use with an un-narrowed excimer,



1.23 Catadioptric projection objection lens using a beam splitting cube to allow for a full pupil imaging without central obscuration. The use of polarized illumination allows for throughput near unity.

but refractive approaches eventually dominated. As 193 nm lithography evolved toward very high NA values, the use of reflective components within projection lenses increased again, leading to new catadioptric lens concepts for leading edge lithography, such as the lens designs depicted in Fig. 1.15.

1.8 Resolution enhancement technology

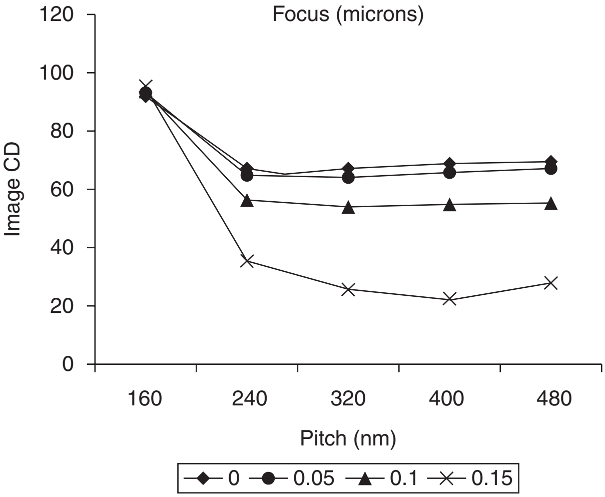
The Rayleigh equation suggests that the limit to resolution for a given wavelength and NA can be achieved using incoherent illumination, where the k_1 factor is 0.25. Although it is true that some degree of image modulation may result from mask geometry at this minimum pitch (accounting for optical reduction factor), the usability of such an image is not certain. It is likely, in fact, that incoherent illumination of features at $k_1 = 0.25$ will result in an image modulation so poor that photoresist can make no use of it even at best focus. The concept of resolution enhancement technology (RET) is a general one used in many fields of imaging. In the case of optical lithography, RET generally refers to methods employed to improve on the quality of an image compared to what could be achieved using incoherent illumination alone. Examples of now commonly used RET include phase shift masking (PSM), optical proximity correction (OPC), modified or off-axis illumination (OAI), and polarized illumination.

1.8.1 Mask correction approaches and optical proximity correction

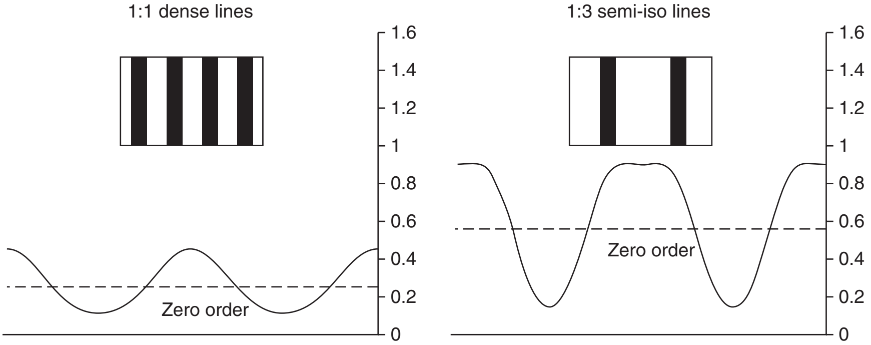
Features on a photomask of the same size may image very differently depending on how they are grouped and what is neighboring them. This is primarily driven by the feature density dependence as seen in the Rayleigh equation, where resolution is determined by the periodicity or pitch of an object. Theoretically, geometry of any size could be imaged, regardless of how small it is, so long as its periodicity falls above diffraction limits. This is why the minimum half-pitch designation of a resolution limit is more useful than the minimum feature size. In practice, however, the resolvability is determined by a photoresist's ability to print such an image into a well-defined three-dimensional structure. Because of the variation in the printing ability of a small feature is strongly determined by what is in its proximity, modifications to its surrounding can influence performance. This is often referred to as OPC, which is used to correct for optical proximity error (OPE). Two types of OPC are often employed at a photomask. One addresses the feature density dependence described above. A second type is used to compensate for feature corner rounding and line shortening effects in the image plane (the X–Y plane).

The plots in Fig. 1.24 show the first type OPE, namely the behavior of a line of a given size through pitch and defocus. The example shows an 80 nm line imaged using a 193 nm wavelength and a 0.80 NA objective lens with a partial coherence (σ) of 0.80. Four focus values are shown, from best focus to 0.15 μm defocus. The image CD is plotted, which is the width of the aerial image at a 30% intensity level. This would correspond to the approximate level of response of a photoresist. As is clear from the figure, the behavior of the 80 nm line is very dependent upon its density. In particular, the image CD decreases as the periodicity decreases. Uniform CD performance is desired to meet device performance requirements.

The reason for the differences in the imaging behavior of features with changing density can be understood by considering the neighborhood surrounding a line as density is varied. Shown in Fig. 1.25 for example are lines with duty ratios (or line-to-space ratios) of 1:1 and 1:3. Also shown are the coherent aerial images for each case when few diffraction orders are collected. The aerial image of 1:1 lines modulates around a background intensity of 0.25, whereas the image for the 1:3 lines modulates around a background value of 0.56. The reason for these differences is fundamental, where background intensity corresponds to the 0th diffraction order squared, or the space-to-pitch ratio squared, $(s/p)^2$. These two situations can have very different printing performance as focus and exposure are varied. The challenge of OPC is therefore to equalize the background levels across a variety of duty ratios.

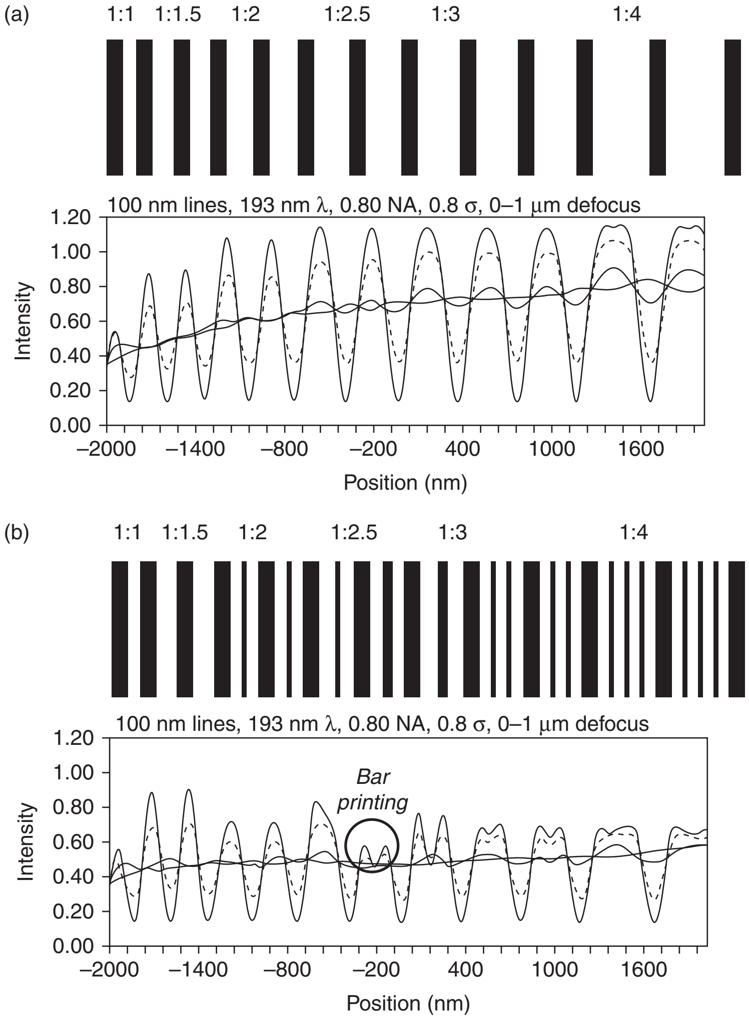


1.24 OPE for lines of various pitch values through-focus, showing how neighboring features influence behavior.



1.25 Aerial image plots for lines of two duty ratios, 1:1 and 1:3, showing how differences in line spacing drive the average image intensity.

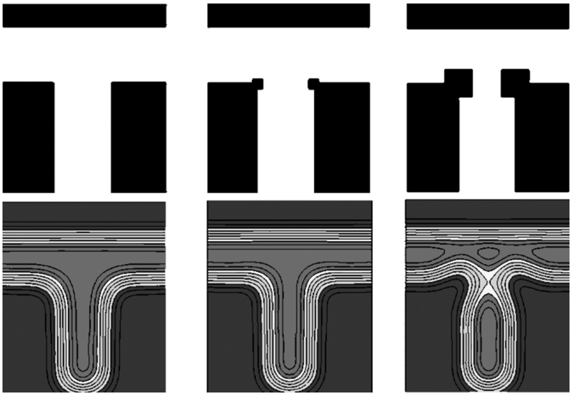
Figure 1.26 shows how the use of additional features to a mask can lead to a more common through-focus behavior. Figure 1.26a is the through-focus intensity of these lines with duty ratio from 1:1 up to 1:4. As defocus is increased, the intensity around which the lines demodulate increases by as much as 2×. By adding correction features as small bars, as seen in Fig. 1.26a, the intensity between the lines with increasing duty ratio can be decreased, leading to a more common through-focus intensity behavior. The goal for the bars is to influence this demodulation behavior but not lead to printing of the bars themselves. Regions where bars may begin to print are those



1.26 Aerial image plots for a mask with 100 nm lines of varying duty ratio with (a) no correction for the OPE effect, and (b) correction features for the OPE effect, where the addition of small features produces a level average intensity through pitch.

where the bar width becomes large compared to the space within which it is placed. The use of multiple thin bars can help to reduce bar printing effects.

The second type of OPE and subsequent OPC is the use of additional mask features (serifs and other shapes), which can alter the energy collected by the projective lens for parts of a mask feature that are near line ends and corners. An example of this is shown in Fig. 1.27, showing a simple T-shaped

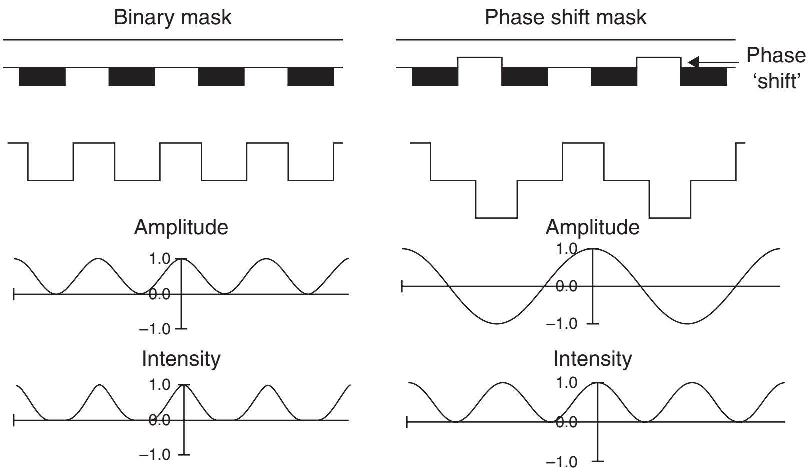


1.27 OPC using small serifs at the corners of features to compensate for the energy losses in these regions resulting from the lack of capture by the objective lens.

feature on a mask together with a plot of the intensity at the image plane when a limited amount of diffraction energy is collected. For imaging conditions where few diffraction orders are collected, rounded corners result as high frequency diffraction orders are lost, and the formation of square corners is thus difficult. The situation can be compensated to some extent by over-exaggerating the corners to allow more exposure energy in these regions. The situation becomes one of dose compensation to help account for the energy that is lost. This is not actually a 'correction' method since the diffraction energy is at frequencies too high to be collected by the projection lens. It is instead an optical proximity compensation approach to allow for printing performance closer to what is desired. Figure 1.27 shows how this image compensation can be carried out using serifs at the corners of rounding features. Too much compensation can cause over exaggeration of the effect as well.

1.8.2 Phase shift masking

As seen earlier, a conventional 'binary' photomask is defined using an amplitude function that varies between values of one and zero. There is potential to design into a mask different values of phase, which can lead to image variation and improvement. Consider for example Fig. 1.28. The amplitude and intensity functions for a binary grating mask are plotted for a limited amount of diffraction energy collected by an objective lens. The amplitude function is the lowest frequency sinusoidal representation of the mask function. When squared, this leads to an intensity function that can be used to expose photoresist. If the grating mask is modified with phase



1.28 Comparison of a binary mask and an alternating PSM, showing the decrease in the image amplitude frequency for the PSM.

shifted regions between every other line, the resulting amplitude function is sinusoidal with one half of the frequency. When squared, the resulting intensity function resembles that of the binary mask (a squared sinusoid vs a squared biased sinusoid). The differences, however, are significant in two ways. First, since the frequency of the amplitude function is 50% lower than the binary amplitude function, the objective lens pupil could be one half as large to collect all of the diffracted energy necessary. Thought of another way, a lens of a given NA could image a grating of twice the frequency for a phase shifted mask compared to a binary mask. Also, since the amplitude function passes through zero with a phase shift, an intensity value of zero in the dark mask regions is ensured even with image degradation (such as defocus).

Achieving a phase shift of 180° (or π) is necessary to bring about such effects. This is accomplished through multiple process steps during the mask fabrication process, etching regions in a mask so that a wavetrain of light propagating through a phase shifted opening is out of phase with an unshifted opening by half a wavelength. The example shown in Fig. 1.28 is for an ideal grating mask, which is simpler than the geometry of real ICs. When phase shift masks are used, a second binary mask is often used in a dual exposure sequence to remove unwanted pattern regions or artifacts. This second mask is referred to as a 'trim mask.' An example of two such masks (a PSM and a trim mask) are depicted in the high density gate designs shown earlier.

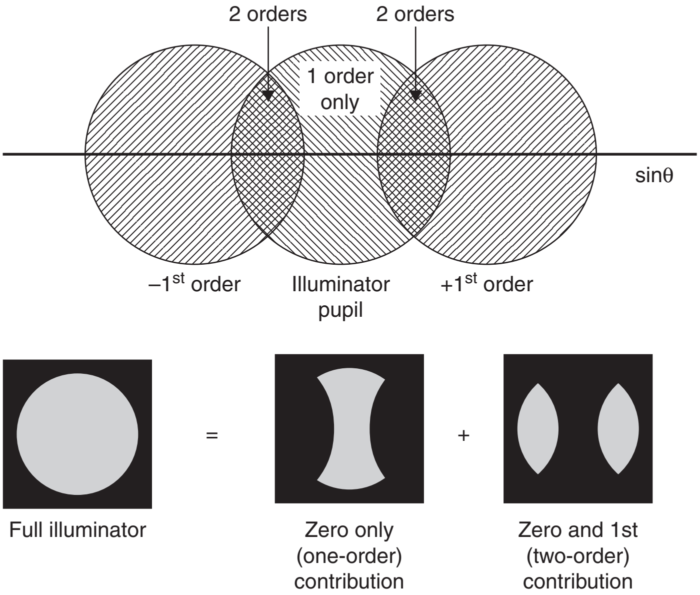
Variations on this basic theme have been developed to adapt PSM to real device applications. These include the use of light 'leaking' through the dark regions of a mask, allowing for a phase shift to clear regions through appropriate selection of materials and film thicknesses. These types of mask are known as attenuated phase shift masks (APSM) and can utilize materials such as molybdenum oxide and silicon dioxide (Mo-Si APSM) or other refractory metal oxides and nitrides combined with silicon dioxide or silicon nitride.[6]

1.8.3 Modified illumination

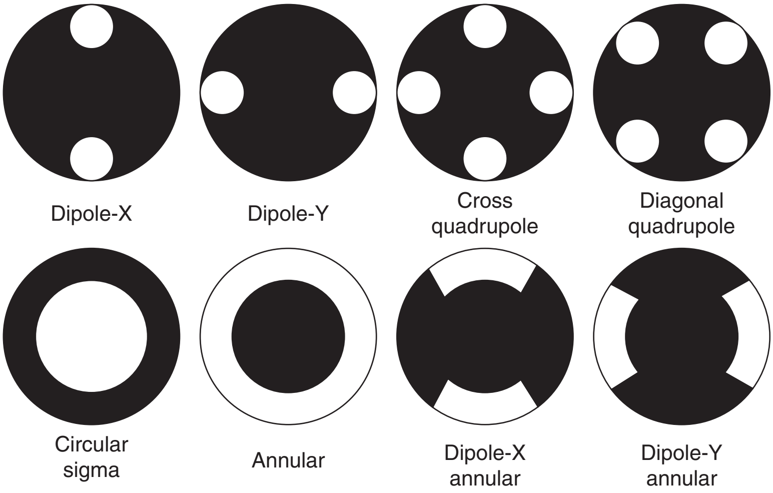
Up to this point, mask illumination that is circular and centered on the optical axis has been considered. Alternatives to this 'on-axis' illumination are also possible. Using OAI, source shaping such as dipole, quadrupole, annular, and more customized variations can lead to improvements in the printing of photoresist features. Although incoherent illumination leads to the limits of resolution for an optical system (and a $k_1 = 0.25$), the image quality at these limits will generally be poor. This is especially true as defocus and other image degradation mechanisms (e.g. aberrations) are considered. For projection optical lithography, incoherent illumination should be thought of as a starting point, where improvements in image quality can be achieved through the removal of select portions of a full $\sigma = 1$ illuminator. This approach leads to an understanding of how OAI and more customized illumination are an integral part of leading edge optical lithography.

Off-axis illumination

Though a common treatment of OAI is to start with a coherent on-axis point source, it can be more intuitive to begin instead with a full circular incoherent source, as shown in Fig. 1.29. For X-oriented features with $0.25 < k_1 < 0.5$, there is a central portion of a full sigma illuminator that will produce diffraction energy that cannot be collected by the objective lens pupil. There are portions toward the edge of the illuminator, however, that will result in first diffraction order energy that can be collected by the objective lens. Because both 0th and 1st order energy are collected, it can be used in image production. The left edge of the illuminator results in collected +1st diffraction order energy and the right edge results in −1st order collected energy. If regions without more than 0th order are discarded, what remains is a two-component source often referred to as a dipole illuminator. If this exercise is carried out in the Y direction as well, a four-pole or *quadrupole* illuminator arises. If all mask orientations are allowed, the illuminator becomes *annular* in shape. Figure 1.30 shows a variety of off-axis source shapes that could be utilized in projection lithography. Variations in the pole shape result from the particular method of implementation and they may be circular, arcs, wedges, or other shapes.

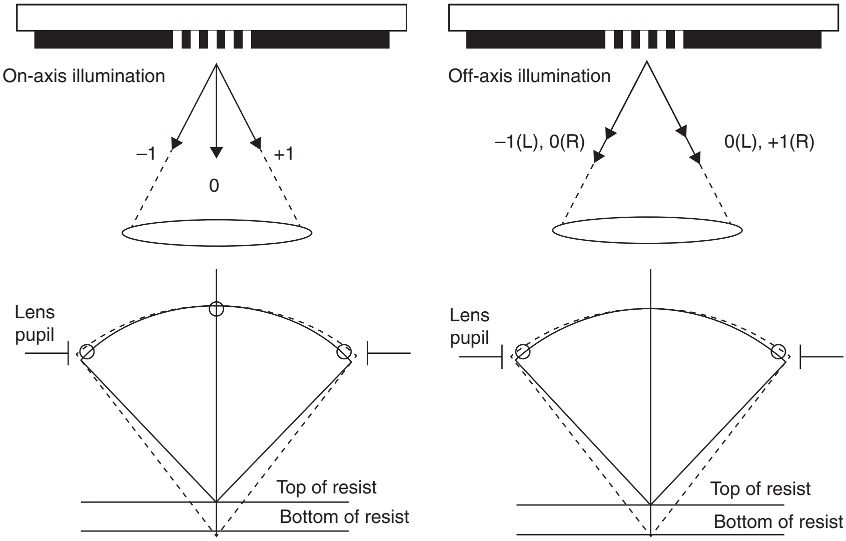

2 orders
2 orders
1 order only
sinθ
−1st order
Illuminator pupil
+1st order
=
+
Full illuminator
Zero only (one-order) contribution
Zero and 1st (two-order) contribution

1.29 Distribution of diffraction energy with OAI and its capture by the objective lens, showing zero and first order contribution to image formation.


Dipole-X
Dipole-Y
Cross quadrupole
Diagonal quadrupole
Circular sigma
Annular
Dipole-X annular
Dipole-Y annular

1.30 Several of the choices for OAI in a projection lithography system.

The advantage in using only a select portion of the lens pupil is appreciated when considering that those portions removed only add background energy to an image. The consequence is a reduction in image modulation.



1.31 Comparison of the capture of diffraction order energy for on-axis illumination and dipole OAI, showing how the off-axis case leads to the avoidance of the center of the lens and a potential for increased DOF.

An additional advantage of OAI is a large DOF that can be gained by distribution diffraction energy only at radial locations of the projection lens pupil. This is shown in Fig. 1.31, where the defocus OPD is shown together with the diffraction order energy for on-axis low sigma illumination and off-axis dipole illumination. In the on-axis case, a large defocus aberration effect will result in image blur because of the phase difference between the diffraction energy traveling through the center of the pupil and that traveling through the edge of the pupil. For dipole illumination, each pole distributes 0th and 1st diffraction order energy at radial pupil locations only. Because of this, there is little difference in the phase of all orders as they are collected by the lens pupil. The end result is imaging with significant improvement in DOF. Indeed, for poles that are single points, the DOF is infinite. More practical source sizes such as those in Fig. 1.30 may be necessary for throughput as well as the reduction in coherent image artifacts, and the improvement will be lessened.

Customized illumination and source–mask optimization

The approach described for source design can be utilized for very specific mask structures to arrive at a unique source shape for the best performance across a mask field (or a critical portion of the field). This can lead to source shaping that has attributes of some of those shown in Figs 1.29 and 1.30, but with additional shaping based on the size, shape, periodicity, and orientation of geometry. This approach can be used together with mask optimization to iteratively optimize the source and the mask through *source–mask optimization* (SMO).[7] Though the concept may be a simple one, the implementation becomes very involved and difficult as large field areas are considered over many billions of device structures. There is additional complexity because of the loss of the phase information in the intensity at the image plane, making an inverse solution to the problem computationally difficult. As phase shift masks are added to the challenges of amplitude (binary) masks, OPC and unrestricted source shaping, the development of efficient approaches and algorithms has been necessary to enable SMO as a practical RET.

1.9 Immersion lithography

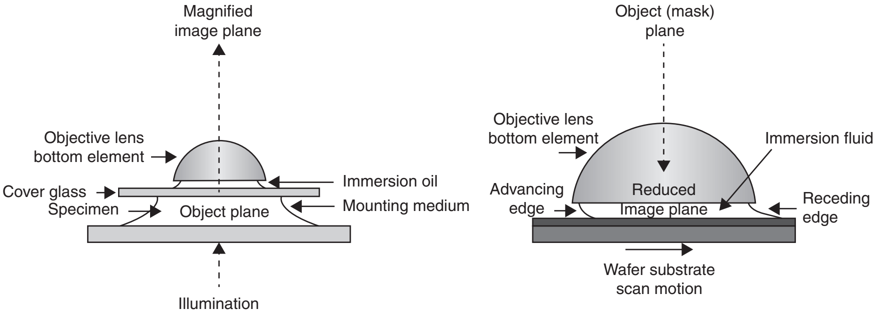
The concept of immersion lithography for high NA had existed for several decades but became practical only when it was apparent that few alternatives existed to extend optical lithography. Additionally, the identification of water as a suitable immersion fluid for use at 193 nm led to the successful implementation of immersion lithography to lithography generations beyond 65 nm.[8,9] The principles of immersion lithography are similar to immersion microscopy, with a few important exceptions. As seen earlier, a lens NA is defined in terms of its collection angle and the refractive index of the surrounding medium:

$$\mathrm{NA} = n \sin\theta$$

The NA is also known as an optical invariant. As light propagates through media of different refractive indices, the NA remains constant, as described by Snell's Law:

$$n_1 \sin\theta_1 = n_2 \sin\theta_2 = n_3 \sin\theta_3 = \mathrm{NA}$$

In an immersion microscope, where magnification of an object at the image plane is desired, the limitation of a lens is its ability to collect large angle diffraction detail from the object. Collection in air is limited to angles of 90° and the corresponding NA of 1.0. If image media of higher refractive index were used, such as immersion oils with indices as high as 2.0, the maximum NA would be increased corresponding to the index, allowing for greater collection of diffracted detail. This would directly equate to higher resolution. An immersion microscope therefore utilizes an increased NA at the object plane through immersion. For optical lithography, projection imaging

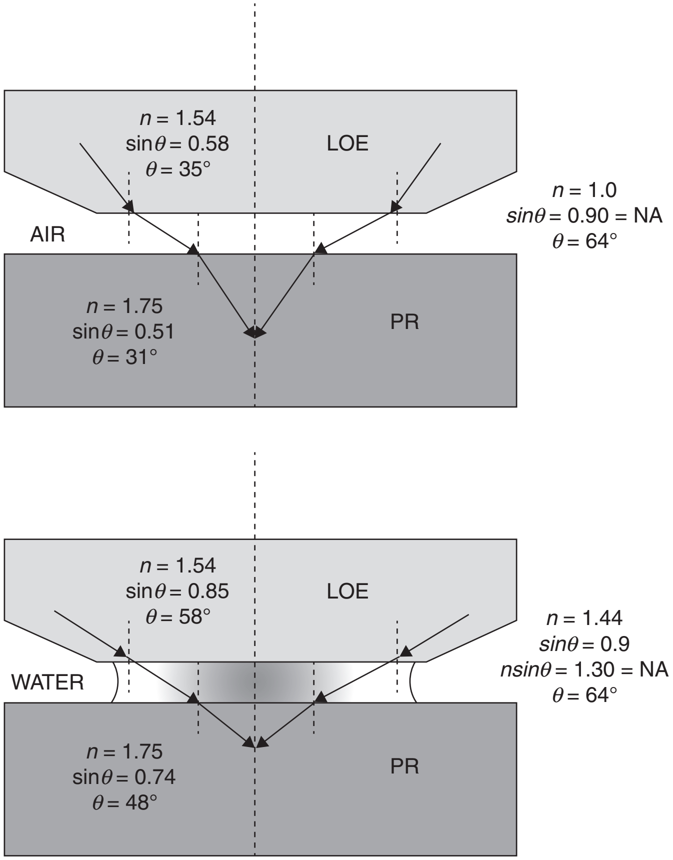

Magnified image plane
Objective lens bottom element
Cover glass
Specimen
Object plane
Immersion oil
Mounting medium
Illumination
Object (mask) plane
Objective lens bottom element
Immersion fluid
Advancing edge
Reduced Image plane
Receding edge
Wafer substrate scan motion

1.32 Comparison of (left) an immersion microscope setup in magnification mode, and (right) a reduction immersion lithography setup.

occurs in a reduction mode. The limits to diffraction are thus imposed at the image plane. Instead of immersing the object plane, as with a microscope, the image plane is introduced to a higher refractive index. The situation is shown in Fig. 1.32.

Another important difference between immersion microscopy and immersion lithography is that, under normal circumstances, microscopy is used with static fields. The meniscus formed at the interface of air, the fluid, and the solids (the cover glass and lens) exhibits a shape that is dependent on these materials properties. In a scanning lithography system, the interface is also unique to the materials used (the photoresist and the lens) and is also influenced by the motion of the stage. A critical requirement of the fluid film is that it follows the surface of the wafer with the lens during a scan. This is facilitated with the water immersion fluid since a lens surface is hydrophilic in nature and a photoresist film is more hydrophobic. The scanning of the meniscus creates advancing and receding edges, as depicted in Fig. 1.32. The angle of these edges must be carefully controlled through the design and engineering of the immersion 'head' containing the optic, the properties of the photoresist material, and the motion of scanning. Inadequate control can lead to defects through the trapping of air into the fluid or the deposition of fluid droplets onto the wafer surface. The control of such defects is critical to ensure successful implementation of an immersion lithography process.

For the DUV wavelengths necessary for leading edge lithography, fluids must be suitably transparent at exposing wavelengths and also be compatible with photoresist materials and processes. Water has very attractive properties in the DUV, with a refractive index that increases from about 1.3 in the visible to about 1.44 at 193 nm. Figure 1.33 shows how water immersion lithography at 193 nm can lead to a higher NA (and therefore higher resolution) than an air (or dry) lithography counterpart. The diagrams compare


n = 1.54
sinθ = 0.58
θ = 35°
LOE
n = 1.0
sinθ = 0.90 = NA
θ = 64°
AIR
n = 1.75
sinθ = 0.51
θ = 31°
PR
n = 1.54
sinθ = 0.85
θ = 58°
LOE
n = 1.44
sinθ = 0.9
nsinθ = 1.30 = NA
θ = 64°
WATER
n = 1.75
sinθ = 0.74
θ = 48°
PR

1.33 Rays traced through the lithography stacks for (top) a dry lithography scenario, and (bottom) an immersion lithography scenario, showing the larger angles enabled with immersion lithography.

situations where the sines of the refracted angles in image media are 0.90. For the dry lithography case, it is clear that the limitation to the refracted angle occurs in the air. For the case of water immersion, the maximum refraction angle in the lens last optical element (LOE), the image media, and the photoresist are closer, and the NA of the system can be increased by a value equivalent to the increase in the refractive index (1.44). NAs up to 1.35 have been achieved using water immersion lithography at 193 nm.[1]

1.10 Multiple patterning optical lithography

Optical lithography options for shorter wavelength and higher NA may have reached a point of maturity. As seen in Fig. 1.18, a wavelength shrink factor below 0.8 would probably be necessary to justify a new generation. This would force optical lithography into the VUV, presenting significant material challenges, as suitably transparent optics, photoresists, mask substrates, pellicles, and immersion media are few and far between. Higher NA for 193 nm immersion is similarly challenging, as a 0.80 shrink through a higher index immersion fluid would require a refractive index of 1.8 or greater. Fluid materials with suitable DUV transmission with necessary mechanical, chemical, and thermal properties are not likely to be identified. In the absence of traditional methods identified in the Rayleigh equation, approaches to multiple patterning at the image plane have been developed. Such approaches could be used to extend optical lithography well beyond $k_1 = 0.25$ and toward resolution that might otherwise be obtained only with other non-optical approaches.[10]

Multiple patterning approaches are also referred to as double patterning, pitch division, and pitch splitting. The concept of these approaches is quite simple. If there is a fundamental pitch limitation for a given process, that limit could be overcome by interleaving two patterning processes with each producing images of small lines (or small spaces) at that fundamental pitch. To appreciate the limitation of the concept, one needs to realize that the interleaved images are ones with minimum contrast – essentially sinusoidal variations in intensity. The interleaving of sinusoidal intensity functions will only lead to constant non-image intensity. In other words, a simple second exposure will wash out a first. What is necessary, therefore, is a method to somehow capture or transfer a first exposure as a high contrast image and follow it up with a subsequent operation to pattern a second exposure. Possible solutions can include a double development method using two development schemes that can discriminate between fully exposed resist and fully unexposed resist and lead to a doubling of resolution through the patterning of feature sidewalls. Alternatively, resist freezing methods can use two photoresist coatings with a fluid or vapor chemical step to freeze the first coating and allow for a second overcoat. As another alternative, a sidewall spacer defined process can lead to the doubling of pitch density through the deposition of a conformal inorganic thin film coating over patterned lines. Subsequent top layer removal steps can then reveal material remaining only on two sides of each feature, resulting in a doubling effect. Or a double pattering approach using a first lithography step, followed by an etch step, followed by a second lithography step, followed by a second etch step (litho-etch-litho-etch) can be used to define lines or spaces. Details of the various processes that have evolved for multiple patterning are described in the resist processing chapter (Chapter 7). Such schemes and variations represent opportunities to extend optical lithography into areas beyond what is possible through wavelength and NA scaling alone.

1.11 Conclusion

Optical projection lithography has been the principle patterning approach for technologies at the microscale to the nanoscale. Advances in optical systems that have included the use of shorter wavelength sources, modified illumination approaches, advanced photomask design, immersion imaging, and multiple patterning have allowed for robust lithography down to a few nanometers. Although it may prove difficult to reduce pattern sizes one hundred-fold again, as has been done over the past few decades, it can be expected that there is more to gain through clever design and insight. Traditional scaling rules continue to be extended into new generations. Combinations of the techniques described in this chapter have not yet been fully exhausted, especially when considered together with the possibilities offered with extreme UV lithography (EUVL, Chapter 2). Also, if hyper-resolution technologies now used in microscopy fields can be applied to lithography, imaging beyond the limits of diffraction may become possible.

1.12 References

1. International Technology Roadmap for Semiconductors (2011), *Executive Summary*, available from: http://www.itrs.net/Links/2011ITRS/2011Chapters/2011ExecSum.pdf.
2. Schenker R., Singh V. and Borodovsky Y. (2010), ‘The role of strong phase shift masks in Intel’s DFM infrastructure development,’ SPIE Design for Manufacturability through Design-Process Integration IV, 7641.
3. Smith B.W. (2012), ‘The saga of lambda: spectral influences throughout lithography generations,’ Proc. SPIE Advances in Resist Materials and Processing Technology XXIX, 8325.
4. Hazelton A., Wakamoto S., Shiraishi K., Fujiwara T., Nagasaka H., Nakano K., Owa S. and Matsuyama T. (2007), ‘Immersion lithography in mass production: latest results for Nikon immersion exposure Tools,’ Sematech Immersion Symposium, http://www.sematech.org/meetings/archives/litho/8065/pres/TO-02%20Hazelton_Nikon%20final%20dist.pdf).
5. Bruning J.H. (2007), ‘Optical lithography: 40 years and holding,’ Proc. SPIE 6520, Optical Microlithography XX, 652004.
6. Smith B.W. (1999), ‘Multi-layered attenuated phase shift mask and a method for making the mask,’ US Patent 5, 939, 227.
7. Rosenbluth A.E., Bukofsky S., Hibbs M., Lai K., Molles A., Singh R. and Wong A. (2001), ‘Optimum mask and source patterns to print a given shape,’ SPIE Vol. **4346**, 486–502.
8. Smith B.W., Kang H., Bourov A., Cropanese F. and Fan Y. (2003), ‘Water immersion optical lithography for 45-nm node,’ Proc. SPIE 5040, Optical Microlithography XVI, 679.

9. Smith B.W., Bourov A., Fan Y., Zavyalova L., Lafferty N. and Cropanese F. (2004), ‘Approaching the numerical aperture of water immersion lithography at 193-nm,’ Proc. SPIE 5377, Optical Microlithography XVII, 273.
10. Maenhoudt M., Versluijs J., Struyf H., Van Olmen J. and Van Hove M. (2005), ‘Double patterning scheme for sub-0.25 k_1 single damascene structures at NA = 0.75, λ = 193 nm,’ Proc. SPIE 5754, Optical Microlithography XVIII, 1508.

2

Extreme ultraviolet (EUV) lithography

B. J. RICE, SEMATECH, USA

DOI: 10.1533/9780857098757.42

Abstract: This chapter describes extreme ultraviolet (EUV) lithography, a photon-based lithography technology with the potential to enable the continuation of Moore's law throughout the 2010s and beyond. In this introduction we motivate the choice of EUV wavelength for use in lithography. A brief history of the early research work on EUV is provided, followed by sections on the phases of development and commercialization. Key EUV technology sub-systems are then covered, with emphasis on the most challenging components: sources, optics, masks, resists and EUV scanner integration. The chapter concludes with ruminations on the economic and technical implications of EUV lithography.

Key words: lithography, extreme ultraviolet (EUV), EUV source, EUV mask, EUV resist.

2.1 Introduction

For the past 45 years the semiconductor industry has followed a trend of technology advancement with clocklike precision. The density of transistors patterned on silicon wafers has doubled every 18–24 months so consistently that the behavior has been dubbed 'Moore's law' after Gordon Moore, the first to observe the trend. This achievement has become more than a marvel of engineering execution – now the steady improvement is an axiom upon which the semiconductor industry's business model relies. The pressure to continue this improvement is more economic than technological.

This chapter describes extreme ultraviolet (EUV) lithography, a photon-based lithography technology with the potential to enable the continuation of Moore's law throughout the 2010s and beyond. In this introduction we motivate the choice of EUV wavelength for use in lithography. We then provide an overview of the key sub-systems of EUV lithography. A brief history of the early research work on EUV is next, followed by sections on the phases of development and commercialization. After the introduction, EUV technology is described system by system, with emphasis on the most challenging components: sources, optics, masks, and resists. Next, some of the integration challenges within the EUV scanner tooling are discussed. The chapter concludes with ruminations on the economic and technical implications of EUV lithography.

2.1.1 Wavelength change

The bread and butter of transistor size reduction, or 'scaling', has been the relentless improvement in optical projection lithography over the past 30 years. Described in detail in the previous chapter, the progression of steadily shorter wavelength lithography steppers and scanners has propelled the industry into a high resolution, high yield space undreamt of in the 1960s when Moore first wrote about transistor densities on integrated circuits. Ultimately, however, physics places limits on photon-based lithographic resolution scaling. The previous chapter introduced the reader to the general form of the Rayleigh criterion:

$$R = k_1\lambda/\mathrm{NA} \qquad [2.1]$$

where R is feature resolution, k_1 is an all-inclusive 'lithography process' proportionality factor, and NA is the numerical aperture of the optical system. For today's optical scanners, the resolution limit of a single exposure is determined by the limits on each of the terms in (Equation [2.1]). The largest NA achievable today on an immersion lithography system is 1.35. The shortest wavelength of light available in a commercial optical lithography exposure system today is 193 nm. The theoretical limit for k_1 in a process using incoherent illumination, a phase-shifted mask and a 'perfectly' tuned photoresist is 0.25. Putting all these together yields a theoretical optical lithography resolution limit for single exposure processes of

$$\frac{0.25\times 193\text{ nm}}{1.35} = 36\text{ nm}$$

Economic forces and worldwide demand for continued improvement in computer chip functionality, however, require that semiconductor device manufacturers fabricate chips with feature sizes much smaller than 36 nm. Referring to Equation [2.1], it is clear there are three approaches to achieving a smaller resolution using photon-based lithography: lower k_1; shorter λ; and increased NA.

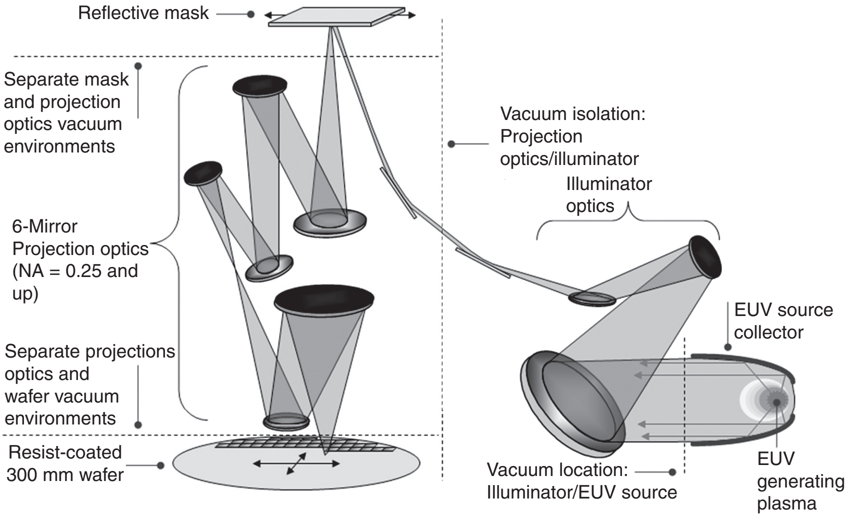
Lowering k_1: it is possible to exceed the theoretical single exposure proportionality limit of $k_1 = 0.25$ by using two or more processing steps in a technique known as double (or multiple) patterning (Lee *et al.*, 2008; Xie and Smith, 2009). For example, 22 nm features can be fabricated in commercial environments using just such approaches. However, the use of twice (or more) the number of processing steps adds cost, making these approaches unattractive.

Increasing NA: for an optical system in air the limit of NA is 1, with a practical engineering limit around 0.93. Today's state-of-the-art optical exposure systems exceed this limit by placing the final optical element of the projection optics in water, effectively boosting the NA by the index of refraction of water ($n = 1.44$) and giving rise to today's NA limit of 1.35. This approach is called immersion lithography. Exceeding the water-based NA limit is possible by replacing the fluid, final lens element, and photoresist materials with so-called 'high index materials.' Such an approach has been investigated, but so far no suitable materials have been found (Zimmerman *et al.*, 2009).

Shortening λ: the final component of the Rayleigh criterion that could allow improvement in resolution is the wavelength. Since photons have no wavelength limit, there is no physics limitation on resolution improvement. In fact, as mentioned at the start of this section, the modern history of lithography is simply a litany of wavelength choices, each shorter than the previous. Extremely bright mercury lamps exhibited strong spectral peaks at 436 nm ('g-line'), 405 nm ('h-line'), and 365 nm ('i-line'), and each of these wavelengths was a broadly utilized commercial lithography tool. As resolution and throughput requirements increased, high brightness mercury lamps were phased out and replaced with excimer lasers. The initial Hg lamp-based 248 nm exposure systems were replaced by krypton fluoride (KrF) laser-based scanners. Argon fluoride (ArF) lasers, emitting at 193 nm, became widely available in the early 2000s and are the predominant technology in use as of this writing, whether in dry or immersion exposure systems. The 157 nm wavelength was explored extensively, but not adopted for commercial use primarily due to the advent of 193 nm immersion lithography just described. X-rays, with a wavelength around 1 nm, were used in X-ray proximity lithography (see Section 2.1.3), but this technology was also not widely adopted.

EUV is a photolithography based on a choice of 13.5 nm for the exposure wavelength. The wavelength choice is 15 times shorter than 193 nm systems, and one that is no longer in the optical band but instead found in the soft X-ray band. This wavelength choice promises the ultimate benefit of vastly improved resolution. Because the improvement is so significant, the usefulness of the technology should extend over a long period of time. This extendibility factor is attractive for semiconductor manufacturers who hope to be able to make use of EUV technology for as long as possible.

It is important to note that there are other ways to improve lithographic resolution that do not involve photons or (Equation [2.1]). Alternative lithography solutions using electron beams, ion beams, and imprint, also


Reflective mask
Separate mask and projection optics vacuum environments
6-Mirror Projection optics (NA = 0.25 and up)
Separate projections optics and wafer vacuum environments
Resist-coated 300 mm wafer
Vacuum isolation: Projection optics/illuminator
Illuminator optics
EUV source collector
EUV generating plasma
Vacuum location: Illuminator/EUV source

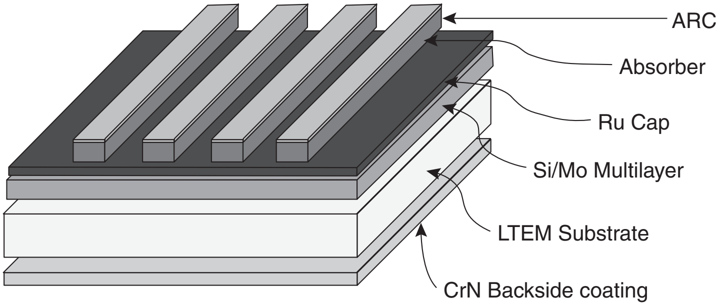
2.1 Key EUV lithography sub-systems.

being explored today, are discussed in their respective chapters elsewhere in this volume.

2.1.2 EUV lithography: the basics

EUV lithography is a projection lithography approach that utilizes 13.5 nm photons to expose photoresist. Figure 2.1 shows a schematic of an EUV lithography exposure system. Before describing the key sub-systems, some general observations are in order. First, since 13.5 nm is in the soft X-ray band, all materials absorb EUV photons strongly, meaning it is not possible to use refractive optics nor is it possible to transmit these photons through air. Second, the reflectivity of essentially all single materials at normal incidence is very small. ‘High’ (e.g. 70%) normal incidence reflectivity can be achieved using a series of paired thin layers of high and low atomic number (Z) materials, such as molybdenum and silicon. Such a construct is known as a Bragg reflector or, more commonly, a multilayer (ML) mirror. High reflectivity can also be achieved using a grazing incident angle.

EUV lithography tools typically utilize a plasma source to generate the 13.5 nm photons. EUV light from the plasma is gathered using an optical element called a ‘collector’. EUV sources and collectors will be described in detail in Section 2.2. Light from the collector is directed into a set of shaping optics collectively known as the ‘illumination optics.’ This light illuminates a photomask located on a high scan-speed vacuum stage. The illumination optics consists of multilayer-coated normal incidence mirrors as well as grazing



2.2 Schematic of an EUV multilayer mask in cross section (not to scale). (*Source*: Frank Goodwin, SEMATECH.)

incidence mirrors. EUV masks (see Fig. 2.2) consist of a six inch square, quarter inch thick low thermal expansion material with a multilayer reflective coating and an absorber layer, typically chrome, etched into the design of a circuit layer. The reflected image of the EUV mask enters a projection optic with a demagnification, typically 4:1. The projection optics are typically a collection of six or more multilayer mirrors and has an NA > 0.25. The final image is focused onto a silicon wafer coated with a photo-sensitive etch resist chemical emulsion or photoresist. The wafer is located on a high scan-speed vacuum-based stage. Each of these sub-systems, from plasma source to wafer stage, operates in a low-hydrocarbon, high vacuum environment.

2.1.3 The EUV lithography research phase

Since their discovery in 1895 by Wilhelm Conrad Roentgen, X-rays of various wavelengths have been used in imaging. Early concepts to shape X-rays using optics were proposed in the 1940s and 1950s by Kirkpatrick and Baez (1948) and Wolter (1952), followed by the first X-ray telescope, conceived in 1960 by Riccardo Giacconi and Bruno Rossi. Later, in the 1970s and 1980s, X-ray proximity lithography (XPL) was proposed, using X-rays and proximity masks. While XPL was successfully demonstrated, manufacturing difficulties with the proximity masks prevented the adoption of this technology.

The principal difficulty with imaging is directing the rays themselves. X-ray machines record images of the contrast between X-ray transmission and absorption of objects placed in the direct line of sight of an X-ray source. In the case of XPL, masks are placed directly under a collimated and condensed X-ray beam (Uda *et al.*, 1998), literally casting a shadow of the circuit pattern etched into the mask onto a photoresist-coated wafer. For X-ray telescopes, grazing incidence optics are used to gently bend the light, permitting the collection of larger solid angle of light and enhancing the imaging capability.

In the 1980s the technology to fabricate normal incidence soft X-ray mirrors was entering a period of rapid improvement spurred by interest in

X-ray lasers for defense, X-ray microscopy, X-ray solar physics and X-ray lithography (Underwood and Barbee, 1981; Zernike and Attwood 1994). Using normal incidence reflective surfaces, it became possible to design more compact soft X-ray projection optics with track lengths on the order of a meter. Normal incidence soft X-ray mirror technology ushered in a new era in lithography research.

Early research on soft X-ray projection lithography was conducted around the globe. The pioneering work was conducted in the mid 1980s by groups at NTT in Japan (Kinoshita *et al.*, 1989), AT&T Bell Labs (Silfvast and Wood, 1988; Bjorkholm *et al.*, 1990; Berreman *et al.*, 1990), various US national laboratories (Hawryluk and Seppala, 1988; Hawryluk *et al.*, 1989; Kubiak *et al.*, 1991), and FOM in the Netherlands (Bijkerk *et al.*, 1991). Together, these groups laid the foundation for the development of imaging with soft X-rays. This research established the feasibility of diffraction limited imaging using soft X-rays and, in so doing, highlighted the technical challenges to commercializing the technology: photoresists capable of high resolution image transfer, sources sufficiently bright to deliver adequate power for high volume throughput, reflective masks with sufficient quality and contrast, and design requirements for optics.

The early research into EUV also defined the wavelength to be used. Throughout the 1980s and early 1990s, soft X-ray imaging research had been conducted over a band from 4 to 40 nm (Kinoshita *et al.*, 1989; Hawryluk and Ceglio, 1993; Stearns *et al.*, 1993). The variety of wavelengths employed was a consequence of material choices in the fabrication of multilayer reflective optics used by the researchers. The multilayers were constructed using pairs of high atomic number (‘reflective’) and low atomic number (‘transmissive’) materials in a quarter wave architecture determined by the Bragg wavelength

$$\lambda = 2d \sin\theta \qquad [2.2]$$

where d is the period of the multilayer structure and θ is the angle of incidence of incoming light rays. The resulting ‘Bragg reflector,’ henceforth referred to as a multilayer stack or multilayer, results in high reflectivity for specific wavelength and material combinations. The ‘best’ material pairs were chosen based on highest net reflectivity but with reasonable wavelength bandwidth. If the bandwidth is too small it does not permit matched transmission of light rays from different angles of incidence emerging from the projection optics.

Another consideration was the requirement that the photons be capable of producing high contrast images in photoresist with reasonable resolution and depth of focus. EUV photoresist performance depends strongly on the

Table 2.1 Optical performance [Ref CXRO], resist performance in PMMA and safety considerations of candidate materials for the multilayer

ML material	λ	$\Delta\lambda/\lambda$	Reflectivity	PMMA profile	Safety
Ni/C	4.4 +	Poor	Poor	Excellent	Benign
La/B_4C	6.7 +	Poor	Good	Excellent	Benign
Mo/Be	11.3 +	Good	Excellent	Good	Toxic
Mo/Si	12.5 +	Excellent	Excellent	Poor	Benign

Source: Kinoshita *et al.*, 1989; Wood *et al.*, 1995; Skulina *et al.*, 1995; Kortright *et al.*, 1993; and Berreman, 1991.

EUV absorbance of the material. Poly(methylmethacrylate), or PMMA, an organic resist representative of EUV imaging materials, exhibits increasing absorption with wavelength from 4.4 nm upwards (Kinoshita *et al.*, 1989). Wood *et al.* (1995) showed that PMMA's sidewall profile performance was best at shorter wavelengths and worsening with increasing wavelength, with excellent behavior at 6.8 nm but considerably degraded performance at 13.9 nm.

Table 2.1 summarizes the characteristics of four different material choices and their associated wavelength ranges. The lower limit of each wavelength range is listed in the table and stems from the X-ray absorption cutoff for the 'transmissive' material in each pair (i.e. C, B, Be, Si). Taking into account the influences of photoresist performance, multilayer reflectivity and bandwidth, and safety considerations arising from materials in the multilayer, the industry settled on a wavelength range just above 12.5 nm as the standard wavelength for soft X-ray projection lithography. Specifically, 13.5 nm was chosen, due to a local peak in xenon photon emission observed in EUV source experiments.

While the industry was converging upon the wavelength choice, the name for the technology was being adjusted. The successes of a decade of research into the feasibility of lithography with soft X-rays had resulted in a general excitement about the prospects of the technology. Since the shorter wavelength X-ray technology XPL was in development and, competing for resources and supporters, the community adopted the term 'extreme ultraviolet' (EUV) lithography to provide easy differentiation between the two X-ray technologies.

2.1.4 The EUV lithography development phase

In the early 1990s, industry and government organizations began to take an interest in supporting the development of the many component technologies that comprise EUV lithography. These efforts generally centered upon evolving the EUV imaging tools from lab experiments to proof of concept systems.

The national EUV lithography program

In 1994, the US Department of Energy (DOE) began the national EUV lithography program. This program funded EUV development efforts at a variety of DOE-sponsored National Laboratories including Sandia National Laboratory (SNL), Lawrence Livermore National Laboratory (LLNL), and Lawrence Berkeley National Laboratory (LBNL). Over a three-year period, progress was made in constructing high resolution EUV imaging equipment, but before the feasibility of a commercial EUV imaging system could be established the US Congress reduced DOE's budget substantially, forcing the DOE to halt all EUV funding.

The EUV limited liability company (LLC)

Recognizing the potential of the technology and the imminent loss of concentrated EUV expertise at the National Labs with the looming end of federal funding, Intel Corporation provided some short-term funding to maintain the EUV research programs. Historically, this was typical behavior for a leading semiconductor company: invest heavily in next-generation technology and secure a competitive advantage through early access to the technology. In the case of EUV lithography, however, Intel rapidly concluded that the cost of perfecting the technology would be significantly more than that of prior optical lithography technologies. A new model was required.

Intel took a visionary approach: create a collaborative model where semiconductor manufacturers would share the cost of developing the technology in exchange for intellectual property and preferential rights to purchase equipment made with that IP. The EUV limited liability company (EUV LLC, or 'LLC') was formed in late 1996. The LLC entered into a Cooperative Research and Development Agreement (CRADA) with 'the Virtual National Laboratory' (VNL), an amalgam of SNL, LLNL, and LBNL, in 1997. This arrangement ensured the seamless transfer of early EUV learning from the previous, government-funded phase to the privately funded LLC.

The LLC program centered on the fabrication of a demonstration tool called the Engineering Test Stand (ETS). The ETS was envisioned to be a state-of-the-art prototype scanning exposure tool that would become the starting point for subsequent commercial EUV scanners. The LLC formed technical programs to support the construction of the ETS: ETS fabrication, EUV optics, ML coating, metrology, masks, EUV source, and resist. These technical programs became the focus of intense academic and industry research and development activities, and effectively created a checklist that could be used to evaluate the readiness of EUV for commercial use. Worldwide efforts expanded with the creation of Japanese (ASET) and European (EUCLIDES, MEDEA, PREUVE) government-funded programs to support development activities being led by the LLC's technical programs.

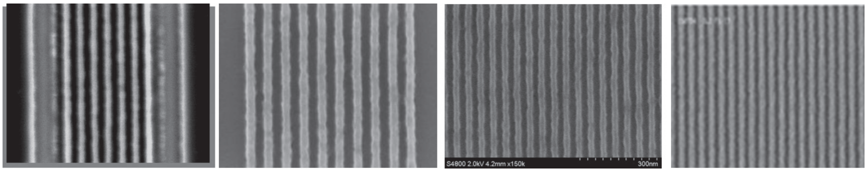
The primary output of the EUV LLC was a working ETS that established the credibility of EUV as a potential commercial lithography, and the identification of key technology challenges that would need to be overcome if EUV was to be implemented in high volume manufacturing. While the LLC effectively established solutions for multilayer coatings and at-wavelength interferometry, in other areas it merely codified the gap(s) that remained.

The consortia

In the early 2000s, the community of EUV researchers was beginning to realize that commercialization of EUV would require vast resources and concerted international collaborative effort. Characterized by a combination of private and government funding, several regional consortia of semiconductor manufacturers began to direct significant resources (in excess of $ 100 M per year) into EUV research. These consortium-based efforts were primarily focused on creating a commercial infrastructure for the implementation of EUV. The SEMATECH consortium, for example, created a number of development centers in the mid 2000s to enable segments of EUV technology.

SEMATECH's Mask Blank Development Center was launched in 2003 with the goal of establishing a commercial supply of EUV mask blanks – ML-coated, defect-free unpatterned mask 'blanks'. The process of fabricating such an EUV blank required access to a number of commercial tools that did not exist. SEMATECH entered into development agreements with a number of equipment manufacturers to develop the required equipment, such as multilayer deposition tooling, mask blank cleaning equipment, and high resolution defect inspection equipment. By 2008, commercial versions of these tools were available for purchase and two mask blank manufacturers were enabled. Over this period of time, EUV mask blank defect levels were reduced from thousands of defects per blank to several tens of defects per blank.

In the photoresist arena, the supplier community required access to reliable imaging equipment to test their formulations. Intel, SEMATECH and SELETE launched programs to create small-field 'micro' exposure tools (MET) that could be used for quick-turn testing by resist suppliers. 0.3 NA METs became available in 2004. By 2010, EUV resist suppliers were evaluating thousands of distinct photoresist formulations per year on the



50 nm, Intel MET, 2004

32.5 nm, SEMATECH LBNL MET 2006

24 nm, SEMATECH LBNL MET, 2009

16 nm, SEMATECH LBNL MET, 2012

2.3 Images of chemically amplified EUV photoresist from the Intel MET (leftmost image only) and SEMATECH LBNL MET highlighting EUV resist resolution improvement since 2004.

SEMATECH METs. Photoresist resolution improved at a steady rate as access to the METs became widespread. Figure 2.3 shows chemically amplified resist resolution improvement over the 2004–12 timeframe. It should be mentioned that 50 nm resolution was first achieved in non-chemically amplified photoresist as early as 1990 (Bjorkholm *et al.*, 1990), but chemically amplified resists did not surpass this until the 0.3 NA METs became available.

2.1.5 The EUV lithography commercial phase

Ultimately, the success of EUV lithography will be measured by its industrial-scale use in the fabrication of semiconductor chips. The chipmakers require that reliable and cost-effective equipment be available for purchase. Because of the inherent complexity of EUV, the cost to develop commercial lithography tools is exceedingly high and requires an entire ecosystem of supporting technology. Commercial suppliers for these support technologies likewise require assurance from chipmakers of the timing and volume of EUV adoption in order to invest. The entire ecosystem investment becomes a series of chicken/egg problems. Investments by consortia, such as those described in the previous section, have helped enable key supporting technology.

In the mid 2000s, ASML and Nikon, world leaders in lithography exposure equipment manufacture, made a large-scale commitment to develop so-called ‘alpha’ EUV exposure tools. Operating by 2006, these tools lent credibility to EUV. The upswing in interest was immediate, and the result was an announcement of a beta tool program by ASML. Six beta tool customers emerged, providing the needed funding for ASML to continue its development program and spurring many of the other suppliers within the EUV ecosystem to increase their independent EUV programs. Today ASML has orders for, and is working to produce, the first production-ready EUV lithography tools, while Nikon is evaluating when to enter the market.

Table 2.2 Specifications for ASML's ADT, NXE:3100, and NXE:3300 platforms

Platform	Alpha demo tool (ADT)	NXE:3100	NXE:3300B*	NXE:3300C*
Resolution	32 nm	27 nm	22 nm	16 nm
NA/σ	0.25/0.5	0.25/0.8	0.33/0.2–0.9	0.33/Off-axis
Overlay (single machine)	<7 nm	<4 nm	<3.5 nm	<3 nm
Throughput (WPH=wafers/hour)	4 WPH	<10 WPH (60 WPH target)	125 WPH target	150 WPH target
Resist Dose	5 mJ/cm^2	10 mJ/cm^2	15 mJ/cm^2	15 mJ/cm^2
Source Power	8 W	10 W achieved (100 W target)	250 W target	350 W target
				* Forthcoming

Source: Chen, 2010.

Table 2.2 shows the top level performance characteristics of ASML's EUV exposure systems to date and the projected capability for those systems that are forthcoming.

Without a mature ecosystem of support infrastructure, components, subcomponents, and materials, however, integrated exposure systems such as those described in Table 2.2 would simply not be achievable. Coincident with ASML's investment, companies have been working to commercialize EUV optics, EUV sources, EUV masks, EUV resists, EUV mask carriers, EUV mask deposition, clean, inspection, review and repair tools, and many other portions of the EUV ecosystem. The total investment in all aspects of EUV has surpassed $5 billion.

2.2 EUV sources

Photolithography is the technique of using light to transfer patterns. It is therefore appropriate to begin the discussion of the component technologies of EUV lithography with an examination of the approaches used to generate light at a wavelength of 13.5 nm. Devices which create these photons are called sources.

2.2.1 EUV source approaches

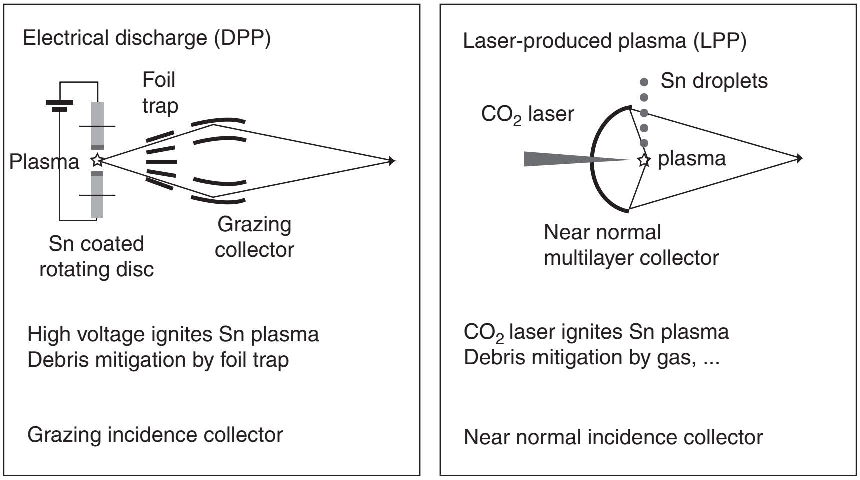
Optical lithography at various visible and ultraviolet wavelengths employs either high power lamps or lasers to provide photons to exposure tools. The power output in the band of interest is in the kilowatt range. The optical lithography exposure equipment generally has a surfeit of photons permitting a high rate of exposure of patterns on silicon wafers limited only by the speed of the wafer and mask stages within the tool. This mode of operation is aptly referred to as *stage-limited*. In contrast, it is the EUV photon source that is the bottleneck in EUV exposure tool throughput. We describe this situation as *photon-limited*.

In the case of EUV lithography, the soft X-ray nature of the 13.5 nm photons prevents adoption of the primary optical methods of photon generation. Lamps and other 'light bulbs' are excluded because of the extremely low transmission of EUV light through material media. For all practical purposes, the 'glass bulb' would absorb all of the EUV light. While low-power (< 1 mW) soft X-ray lasers have been demonstrated, high-power soft X-ray lasers are generally unachievable because of the low reflectivity of even the best multilayer mirrors at 13.5 nm. Recall that laser cavities achieve gain by stimulating photon emission over many passes of a photon back and forth between highly reflective mirrors. With peak reflectivities in the 70% range at 13.5 nm, there is no opportunity for gain and hence little chance to generate a high-power laser output at the EUV wavelength.

As an aside, if one does not require the mirrors to last, it is possible to achieve a short-lived X-ray laser pulse by starting with an ultra-intense burst of incoherent X-rays. Only a few passes in a laser cavity would be required to generate a coherent pulse, albeit with greatly reduced intensity and at the cost of destroying the X-ray reflectors. This is the underlying principle of defense-oriented X-ray lasers.

In practice, two approaches have the potential to produce sustained high-power levels of EUV photons: synchrotrons, and high temperature plasmas. Synchrotrons bend electrons using magnetic fields in various configurations to generate X-ray photons. With circumferences in the 100–1000 m range, installing synchrotrons at modern semiconductor manufacturing plants would pose challenging space and cost obstacles. Smaller synchrotrons, with circumferences of a few meters, can be used in combination with a resonant laser cavity (Loewen, 2003) to produce X-rays. This technique is presently capable of generating only a fraction of a watt of EUV power, and so is currently insufficient to meet EUV exposure tool requirements.

Compact (i.e. room-sized or smaller) plasma generation devices, or plasma sources, have been demonstrated using a variety of techniques. Figure 2.4 shows the two most common approaches: discharge produced plasma (DPP) and laser produced plasma (LPP). LPP and DPP sources are under development as the most viable approaches to achieving reliable high EUV power sources. As of writing this, that potential remains unrealized due to a number of complex manufacturing issues that require mutually compatible


Electrical discharge (DPP)
Foil trap
Plasma
Grazing collector
Sn coated rotating disc
High voltage ignites Sn plasma
Debris mitigation by foil trap
Grazing incidence collector
Laser-produced plasma (LPP)
Sn droplets
CO_2 laser
plasma
Near normal multilayer collector
CO_2 laser ignites Sn plasma
Debris mitigation by gas, ...
Near normal incidence collector

2.4 Typical configurations for LPP and DPP EUV sources. (*Source*: ASML.)

solutions. The balance of this section discusses the source components and their attendant challenges.

Discharge produced plasma sources

DPP sources create a pulsed plasma by periodically arcing a high current through a fuel-filled gap between two electrodes. The electric energy ionizes the atoms of the gas and creates various excitation states of the remaining electrons. The electron states then transition back to lower energy states by emitting photons.

Laser produced plasma sources

LPP sources also create pulsed plasmas, but typically do so using a pulsed laser burst to irradiate a liquid or solid piece of fuel. The fuel is often a droplet that is 'falling' and must be accurately 'hit' by the laser pulse. Once irradiated, the droplet is ionized into a plasma with appropriately populated excited atomic states that then de-excite and emit EUV radiation.

2.2.2 Source fuels

The physics of EUV plasmas has been described by Hassanein *et al.* (2005), but the defining characteristics of the plasma are strongly influenced by the choice of plasma material, or 'fuel,' such as the required input energy, and the relative intensity of the emitted spectrum of photon energies according to the fundamental atomic properties of the material. The goal is to efficiently convert the input energy (electric energy for DPP, laser energy for LPP) into EUV photons (13.5 nm ±1%), the so-called 'in-band' radiation. Unfortunately, most materials that radiate appreciably near 13.5 nm do so in a relatively 'broad' rather than 'peaked' fashion. The net result is that most of the photons generated from the plasma are emitted far from 13.5 nm, in the form of 'out-of-band' (OOB) radiation.

A number of potential fuels have been utilized since stand-alone (that is, not connected to a synchrotron) EUV imaging systems began to be constructed. Investigations of most of the periodic table took place historically throughout the age of atomic spectroscopy, and researchers at the VNL and in industry benefited from that knowledge. Xenon (Xe) was found to have approximately 1–2% conversion efficiency from input power to in-band radiation. As a noble gas it is very attractive because it tends not to react with any of the other source components. Tin (Sn) was observed to have a better conversion efficiency (up to 3–5%), although it tends to deposit and accumulate rapidly on source component surfaces.

Commercial DPP and LPP sources using both Xe and Sn have been constructed and attached to EUV imaging systems. Because of its higher efficiency, all the high volume commercial sources currently under development use Sn exclusively.

2.2.3 The collector

In both DPP and LPP sources, EUV photons from the plasma are emitted in all directions. In order to be useful to an exposure tool they must be directed toward the imaging optics in an organized fashion. An initial optical element called a 'collector' gathers a portion of the photons and focuses them at a point called the 'intermediate focus.' When EUV source power is referenced for a given source it is generally the power measured at this intermediate focus.

DPP and LPP plasmas tend to have distinct shapes. This can have significant consequences on the amount of collectible light. The object and image in a precision optical system are related by the 2D invariant known as étendue, or

$$E = \pi A\,(\mathrm{NA})^2 \tag{2.3}$$

where A is the area illuminated by the optics with a numerical aperture NA. This quantity describes the light gathering power of an optical system. Any photon originating outside the étendue cannot be transmitted through the optical system, and is effectively wasted. For this reason, EUV optics designers tend to favor sources with a small plasma volume capable of fitting entirely within the étendue of the optical system. In general, DPP plasma volumes are larger (~1 mm radius) than LPP plasma volumes (<0.1 mm radius).

DPP sources require quite a close proximity between the electrodes and the plasma, often exhibiting gaps between electrodes of a few centimeters or less. As shown in the right portion of Fig. 2.4, this effectively restricts the usable photons to those that are moving away from the circuit equipment. Because of this constraint, DPP source collector geometry is typically a series of concentric, grazing incidence, and cylindrically symmetric mirrors called 'shells.' Multi-shell collectors can collect as much as 30% of the light emitted from DPPs.

LPP sources are not restricted by the need for electrodes to be near the plasma, and so may use a normal incidence mirror. This collector, like any other normal incidence EUV mirror, must be coated with a reflecting multilayer coating. Compared to a multi-shell grazing incidence collector, this geometry is elegant, simple, and more efficient, theoretically collecting nearly 50% of the light from the laser plasma. Five steradian collectors, with 40% collection efficiency, have been demonstrated.

2.2.4 Debris mitigation

The plasma process is highly energetic (several Joules per pulse), high frequency (kilohertz), and short-lived (hundreds of nanoseconds). The result is that many charged ions, neutrals, and larger clusters, collectively referred to as 'debris,' are created with kinetic energies measured in the keV range. The interactions of the debris with the collector, in particular, can have disastrous consequences. Energetic particles striking the collector mirrors would rapidly sputter the surfaces, effectively ruining their ability to reflect EUV photons. Species such as Sn tend to deposit on mirror surfaces and rapidly reduce the reflective properties of either a grazing incidence or normal incidence mirror. In still other conditions, the presence of the plasma can lead to chemical changes in the collector surface, such as oxidation, also leading to a rapid degradation of reflective properties. These various mechanisms necessitate a debris mitigation effort.

Any successful debris mitigation scheme must do two things: maximize the transmission of EUV photons, and maximize the life of the collector mirrors at high reflectivity. Some of the techniques employed include:

Buffer gas: placing either a background pressure throughout the chamber or a high pressure 'curtain' of gas between the plasma and collector. Debris interacts with the buffer gas and is slowed or deflected, effectively protecting the collector.

Foil trap: for grazing incidence collectors, using a fixed or rotating radial 'fin' configuration that allows photons to pass with high geometrical efficiency (>80%) but has a 'long' longitudinal extent. The principle is that atoms, or more commonly clusters of atoms, moving toward the collector will not typically follow 'straight line' trajectories. Each 'bounce' from a fin or 'foil' will generate an opportunity to stick to the foil, effectively trapping the debris (Corthout and Neff, 2007).

Rotating foil trap: adding a rotation to the assembly (Corthout and Neff, 2007) will increase the number of bounces for the debris dramatically while having no effect on the photons.

Magnetic fields: in the case of charged debris, magnetic fields can be used to deflect ions away from the collector toward a designated dump area (Fujimoto *et al.*, 2012).

Secondary plasma: a secondary plasma (in addition to the one generating the EUV photons) can be employed to ionize neutrals, permitting neutral species to be deflected by magnetic fields as well (Jurczyk *et al.*, 2005).

One technique, complementary to debris mitigation, is debris prevention. In many source fuel delivery schemes, the plasma is formed using less than all of the fuel atoms present. Those excess fuel atoms do not contribute to light generation but do contribute to debris generation. Using so-called 'mass limited targets' reduces the excess fuel leading to correspondingly less debris at the same light output (Richardson *et al.*, 1998).

Neither of these techniques (mitigation and prevention) is 100% effective. In each case, some method of cleaning is necessary due to the high cost of fabricating the collector mirrors. Reactive gas cleaning is a technique employed to remove deposited material from the collector mirror surfaces *in situ* (Rettig *et al.*, 2005). Collector optics lifetimes range from one to three months and correspond to less than 10 000 wafer level exposures (assuming 300 mm wafers) before collector replacement/refurbishment is required. This must be improved by at least a factor of ten to meet HVM requirements.

2.2.5 Source scaling

EUV sources must generate extremely high in-band power, with high reliability and low cost. Cost-effective exposure tools require between 100 W and 250 W of EUV in-band power at intermediate focus. This is an extremely formidable challenge. In the case of DPP, 20 kW or more energy must be dissipated and literally liters of tin must be prevented from coating optical surfaces. For LPP, sufficiently powerful drive lasers and high repetition droplet generators must be fabricated to operate reliably at high duty cycle. Today's achieved EUV source power (about 10 W) is far from the level required (approximately 100 W) to generate commercially viable (70–80 wafers per hour throughput) EUV exposure systems.

2.3 EUV optics

Following the source, EUV photons emerge from the collector to enter the optics of the imaging equipment. Recalling Fig. 2.1, the light from the source is first shaped in the illumination optics and is then directed through the projection optics.

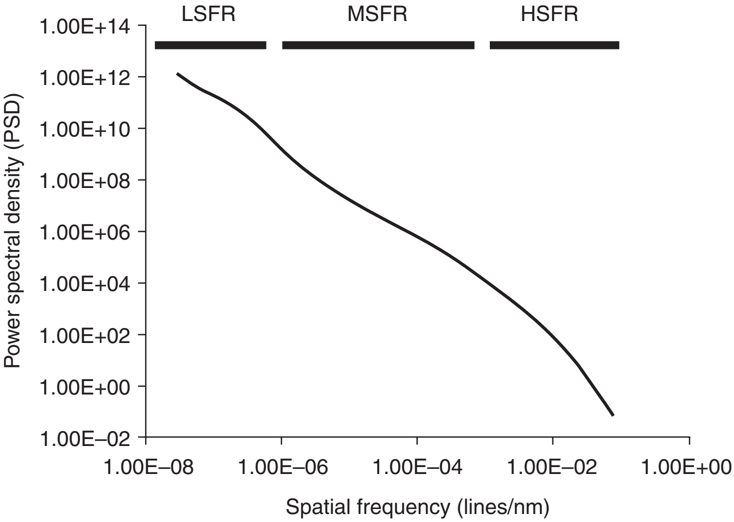
As mentioned in the introductory section to this chapter, normal incidence EUV optics relies on a multilayer stack of thin films to create a Bragg reflector. These mirrors are constructed using specially fabricated low thermal expansion materials (LTEM) because the consequence of 70% reflective mirrors is 30% absorbing mirrors. Even at high heat loads the mirrors must retain their shape, or risk deforming to the point that they no longer produce high fidelity images. The LTEMs have a coefficient of thermal expansion below 10^{-6}.

This section begins with a discussion of the surface requirements for EUV mirrors. Techniques for depositing multilayer stacks are then reviewed. The requirements on the projection optics are described next. This is followed by an overview of issues related to the lifetime and contamination of EUV optics (other than the collector). The section concludes with an investigation of the limits of EUV optics and what extensions are possible.

2.3.1 Surface roughness

While EUV's greatest asset is resolution enhancement due to its wavelength, this very attribute results in a number of considerably tighter specifications. Specifically, light that is incident on a surface has a component that is scattered. Unfortunately, the amplitude of the scattered light is proportional to $(\sigma_s/\lambda)^2$ where σ_s is the rms surface roughness. For comparison, EUV scatters about 200 times more than 193 nm light would scatter from the same quality surface. The implication, therefore, is that surface roughness must improve dramatically to compensate for the increased propensity for EUV photons to scatter.

The scatter from an optical surface is frequently described using the power spectral density (PSD). Figure 2.5, a PSD from a typical EUV mirror, shows the relative strength of signals from grating structures at various length scales, from a few nanometers (lower right) to a centimeter (upper left). The main value of a PSD plot is to identify polishing problems: a smooth PSD curve indicates high quality polish, while a jagged PSD response indicates polishing artifacts at that length scale. The response is broadly classified into



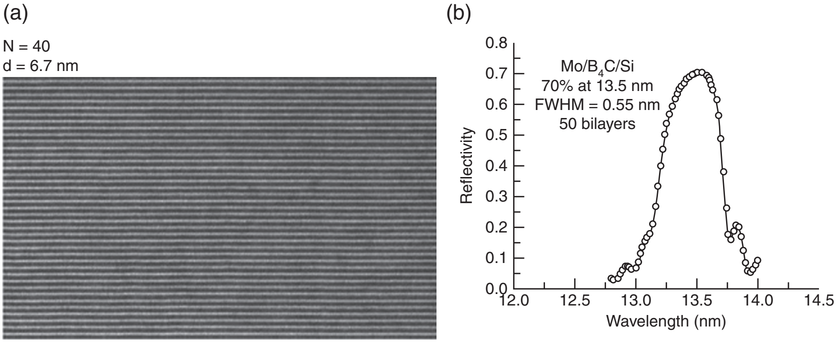
2.5 PSD in arbitrary units of surface roughness on the SEMATECH 0.3 NA MET optic at LBNL. (*Source*: Erik Gullikson, LBNL.)

three categories: low spatial frequency response (LSFR), or 'figure,' above 1 mm; mid spatial frequency response (MSFR), or flare, from about 1 μm to 1 mm; and high spatial frequency response (HSFR) from about 1 μm down to about 10 nm. LSFR represents light that is modestly deflected due to variations over very large length scales. This can impact the fidelity of imaging and induce optical aberrations such as astigmatism. MSFR indicates mid-range deflections of light that provide a background light to other features in the image, effectively reducing the imaging contrast at the wafer. Finally, HSFR corresponds to light that is deflected at very large angles completely out of the image, effectively indicating light loss.

To reduce the flare and aberrations of an EUV optic, it is therefore necessary to achieve better surface roughness at each of the length scales corresponding to the frequency response ranges above. In practice, for EUV mirrors to achieve flare values comparable to 193 nm equipment (about 4% or better), it is necessary to polish surfaces to roughness less than 0.1 nm rms in the MSFR ranges (Litt *et al.*, 2002; Kuerz *et al.*, 2008; Feldmann *et al.*, 2010). This is, on average, less than one atomic layer of variation!

2.3.2 Multilayer deposition

EUV mirrors attain reflectivity through the deposition of many (usually more than 40) pairs of thin films, each a few nanometers thick. As mentioned in Section 2.1.3, the EUV community settled on molybdenum and



2.6 (a) A Mo/Si multilayer film stack. (b) The reflectivity of a Mo/Si film stack with B4C interdiffusion barrier. (*Source*: Sasa Bajt, LLNL.)

silicon (Mo/Si) as the multilayer materials for EUV mirrors. Figure 2.6 shows a multilayer Mo/Si film stack with 50 pairs. The theoretical maximum EUV reflectivity for Mo/Si mirrors is approximately 74% (Henke *et al.*, 1993; Spiller, 1994). In practice, however, physical properties such as roughness, thickness uniformity, interdiffusion of atomic species, oxidation, and thermal stability conspire to limit the reflectivity that can be achieved to about 70%.

ML deposition can be conducted using a variety of available technologies: physical vapor deposition, magnetron sputtering, ion beam deposition, electron beam evaporation, atomic layer deposition, molecular beam epitaxy, and others. Since the 1980s, magnetron sputtering has been a commonly used technique that has achieved highly reflective surfaces. The technique's popularity is due to its ability to coat large surfaces (300 mm diameter and larger), low cost, support for multiple targets within a single chamber, repeatability, and relatively high deposition rate. One drawback of the technique is that it is relatively dirty, creating large numbers of small particles on the surfaces, which makes it unsuitable for use in coating EUV masks (discussed in detail in Section 2.4).

The primary constraints on the multilayer are maximum reflectivity and lifetime. The transmitted light for the entire optical train in an EUV imaging system is proportional to the reflectivity of the EUV mirrors to the nth power, where n is the number of reflective elements. Small improvements in mirror reflectivity can have a big impact on overall system throughput in such cases. For example, for a system with $n = 11$, increasing reflectivity from 66% to 68% results in a 39% increase in system light transmission. Conversely, a reduction in reflectivity from 66% to 64% for each mirror would incur a 29% loss of light. It is important to recognize that each mirror acts as a band pass filter, so the peak reflectivity of each mirror must also be precisely matched to avoid additional losses in the system.

2.3.3 Optical train

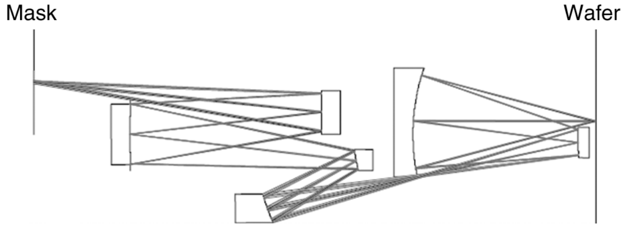
EUV optics are the most advanced optical surface being fabricated today. As mentioned in Section 2.3.1, the surfaces are smoothed to less than an angstrom across spatial length scales that span eight orders of magnitude. In other words, the mirror surfaces are atomically smooth across their entire extent – in some cases hundreds of millimeters! The as-designed surfaces are frequently aspherical, and the actual achieved surfaces deviate from design (the so-called wavefront error or WFE) by less than 1 nm at all points.

Although a detailed description of the requirements for the EUV exposure system's optics is beyond the scope of this text, a few key objectives should be described. Referring to the cartoon in Fig. 2.1, light from the EUV source is first carefully shaped so that upon reaching the mask it is highly uniform across the pupil. This is the function of the illumination optics, or illuminator. The reflected image of the mask must then be demagnified and projected with high fidelity to the wafer. This is the function of the projection optics.

Figure 2.7 shows a ray diagram of a typical projection optic with an NA of 0.25. There are six mirrors in the design, each with varying degrees of curvature. One of the challenges specific to EUV optics is the need for the centroid-reflected wavelength to be constant across all of the points on all of the mirror surfaces. Because the incident angles of light striking the mirror vary with location, the coatings must vary in thickness accordingly to achieve uniform reflectivity across the entire mirror surface. These coatings are called graded coatings.

2.3.4 Optics lifetime

Ultimately, EUV optics must maintain their reflectivity for the life of the exposure system. Depending on the rate of use, this corresponds to a lifetime of four or more years, or about 30 000 h. The lifetime of the EUV mirrors depends on maintaining a minimum uniform reflectivity. Chemical and



2.7 A typical six-mirror EUV projection optic design. (*Source*: Lowisch *et al.,* 2007.)

physical processes and thermal cycling can degrade the ML over time and use, leading to degradation in total reflectivity or uniformity.

The ML stack must first be stable. Mixing of molybdenum and silicon can smear the interface between adjacent layers and, if severe, can impact the constructive interference of the Bragg reflector, reducing overall reflected EUV light. One approach to preventing this behavior is to place a very thin (0.4 nm or less) interdiffusion barrier layer between the Mo and Si layers. Boron carbide (B_4C) is a material typically used for this purpose (Bajt *et al.*, 2002).

The top surface of the ML is susceptible to a number of degradation mechanisms. Chemical processes can cause the top surface to react and form unwanted film growth. Residual water in the EUV vacuum system interacts with EUV light and produces oxygen that then leads to oxidation of the ML. A thin capping layer is used to prevent oxidation of the ML. Ruthenium and titanium oxide are often used for this purpose. Oxidation of these materials self-terminates after the formation of a very thin layer that causes minimal EUV transmission loss.

Contamination in the form of hydrocarbons is an almost unavoidable consequence in an EUV system. A background pressure of hydrocarbons exists in any vacuum environment. EUV photons have sufficient energy to break down or ‘crack’ the hydrocarbons, producing carbon that then deposits on the EUV mirror surfaces. This is highly problematic as even a very thin film of carbon can drastically reduce EUV reflectivity, especially on grazing incidence mirrors where the oblique angles of incidence lead to very long path lengths through the carbon contamination layer. An effective countermeasure for carbon deposition is to use a reactive gas to ‘clean’ away residual carbon. Often a background pressure of atomic hydrogen or oxygen is used to create a steady state of competing deposition and reduction reactions that yields a stable, clean surface. The choice of capping layer must also recognize the need to be impermeable and stable in the presence of such cleaning chemistries as well.

Provided that the strict limits placed on hydrocarbon backgrounds are met continuously during operation, the lifetime of EUV optics meets the required 30 000 h specification.

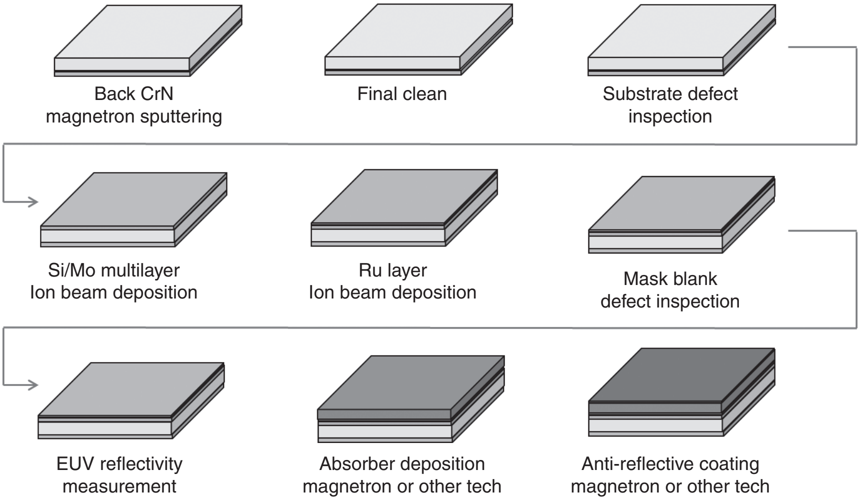
2.4 EUV masks

As mentioned in the introduction, an EUV mask is a 6 inch square, 0.25 inch thick LTEM substrate coated with an EUV reflective ML and covered in an absorbing material etched into a circuit pattern. Figure 2.2 shows a schematic of a typical EUV mask structure. The patterned mask, or reticle, must be a high fidelity representation of the as-designed circuit pattern, or any computer chip imprinted with the pattern will fail to operate correctly. The fabrication process must therefore be defect free, or provide a mechanism to identify and repair any defects.

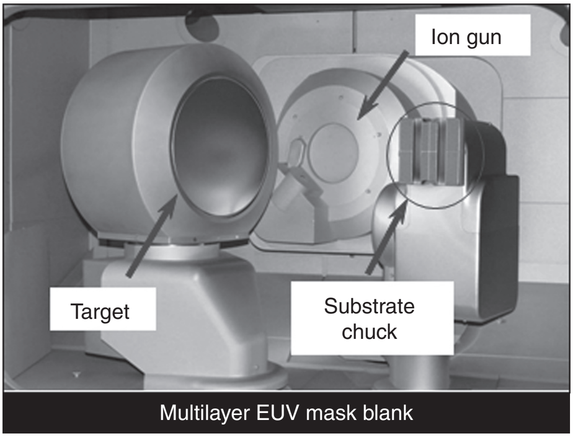
EUV masks differ from optical masks in several important ways. First, like all EUV optical elements, the EUV mask is reflective, while optical masks are transmissive. Second, light incident on the mask is centered on a chief angle that is not normally incident (typically 6° from normal), whereas optical masks receive light centered on the normal. This gives rise to asymmetrical illumination or 'shadowing' of features perpendicular to the light plane. Third, optical masks have a protective, transmissive cover called a pellicle, which is suspended about 6 mm above the etched circuit pattern. EUV masks do not have this pellicle, due in part to the difficulty of engineering a highly transmissive (i.e. < 100 nm thick) membrane with enough structural strength to span the 6 inch square surface of the mask. The pellicle's purpose it to stop particles from landing on the mask surface, holding them at a plane that is out of focus and preventing them from imaging on the wafer. The absence of the pellicle gives rise to a requirement that no particles should be permitted to reach the mask surface.

2.4.1 EUV mask fabrication

EUV mask fabrication is a complex, multi-step process. Figure 2.8 shows this process in block diagram form. The process begins with an LTEM substrate. This substrate must be polished with a peak-to-valley surface flatness no greater than 50 nm.


Back CrN
magnetron sputtering
Final clean
Substrate defect
inspection
Si/Mo multilayer
Ion beam deposition
Ru layer
Ion beam deposition
Mask blank
defect inspection
EUV reflectivity
measurement
Absorber deposition
magnetron or other tech
Anti-reflective coating
magnetron or other tech

2.8 Schematic of the EUV mask fabrication process.



2.9 An IBD tool. Ions emitted from the ion gun strike a target, sputtering target atoms or clusters that uniformly coat an LTEM substrate affixed to the substrate chuck.

A backside metallic coating, typically chromium nitride (CrN), is deposited to permit the use of electrostatic chucks (ESC) in the exposure equipment. The primary reason for the use of ESCs is to eliminate the need to physically touch the patterned side of the mask with a clamp. Physical contact with the surface inevitably generates particles, which in turn are potential defects.

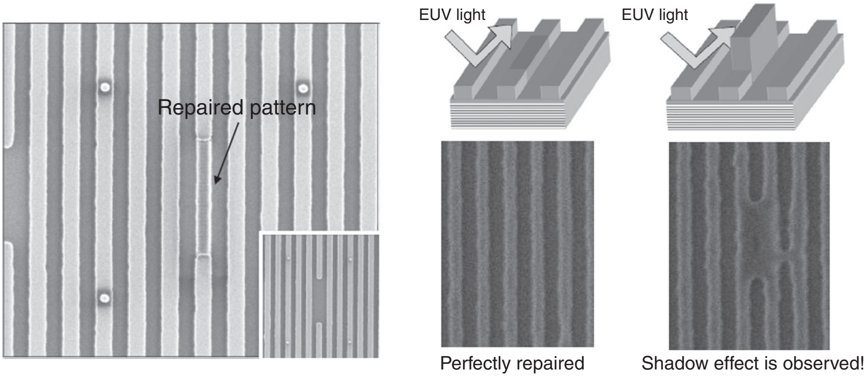
After backside deposition, the mask is cleaned and the front side is inspected for defects in the substrate polishing process. Once the substrate is cleaned and inspected, it is coated with a multilayer. A multilayer-coated substrate is referred to as an EUV mask blank. The equipment used to make mask blank multilayer coatings is different from that used to coat other EUV optics, because the need for cleanliness is paramount. Ion beam deposition (IBD) is the technology of choice. An IBD tool (see Fig. 2.9) directs a beam of argon ions alternately at targets of molybdenum and silicon, physically sputtering atoms and clusters of atoms that are then deposited on an LTEM substrate. The process must be carefully controlled to achieve the required surface roughness and to avoid the generation of particles that could become embedded in the multilayer stack. A capping layer of ruthenium is deposited in the same IBD tool after the formation of the ML stack. The newly created mask blank is then inspected for defects and its reflectivity measured.

The final step in the formation of an EUV mask blank is to coat it with an absorber. This material must not reflect EUV light, instead absorbing it. Since EUV light is strongly absorbed by most materials, this seems trivial. However, as mentioned in the introduction to this section, the light incident on the EUV mask is not normal. The resultant shadowing can result in a biased transfer of patterns. One way to combat this is to make the absorber as thin as possible, thereby limiting the impact of mask shadowing. Another approach is to compensate for the shadowing by resizing features on the mask. Absorber materials and compensation strategies are now the subject of intenive and highly proprietary research.

After the ML stack is finished, antireflective coatings (ARC) are sometimes added to improve the contrast for various inspection tools. The blank is shipped to a mask shop and the absorber is then patterned with photoresist, electron beam writing, and etchers in a fashion similar to that employed for optical masks. An aerial imaging tool that simulates the mask performance in an exposure tool is used to review the patterned mask for defects. Any defects revealed by this technique are repaired, if possible, using local techniques such as focused ion beam etching and deposition. A post-repair review of the aerial imaging tool either certifies the mask for use in production, or flags it for additional repair or to be discarded.

2.4.2 EUV mask defects

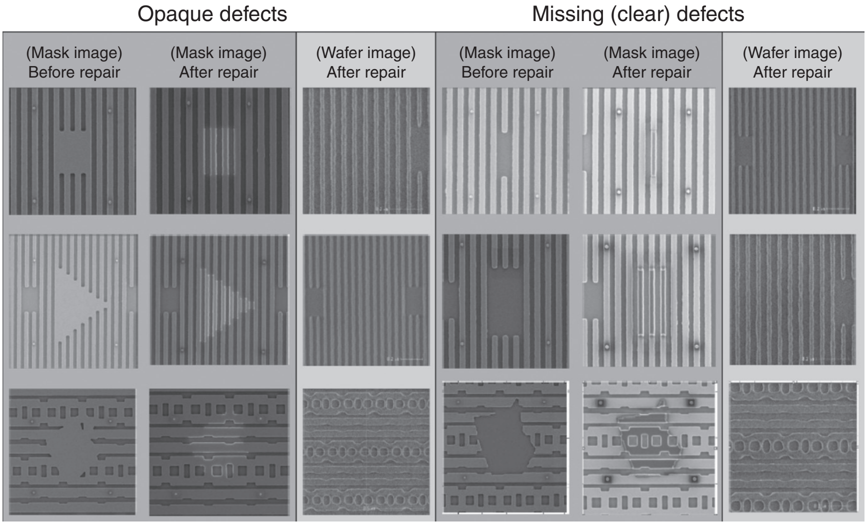
Semiconductor manufacturers require that masks faithfully reproduce their circuit designs without introducing defects that could cause the circuit to fail to perform. For this reason, EUV masks must be fabricated with a minimum of defects. Figure 2.10 shows the variety of defects that can be formed during the mask fabrication process. Defects of interest depend on the size of features being printed with the mask. Any imperfection that would cause a short or an open circuit in the final printed circuitry must be identified



Mask SEM image

Wafer printing results

2.10 Types of defects on EUV masks. (Courtesy Abbas Rastegar, SEMATECH.)

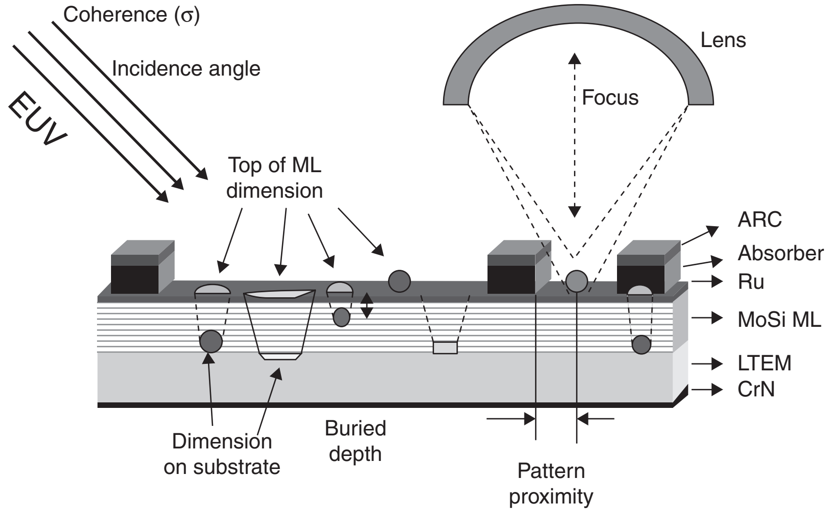

Opaque defects
Missing (clear) defects
(Mask image) Before repair
(Mask image) After repair
(Wafer image) After repair
(Mask image) Before repair
(Mask image) After repair
(Wafer image) After repair

2.11 A typical particle (left) and pit (right) defect on an LTEM substrate.

and resolved (i.e. repaired, ignored due to redundant circuitry, covered with absorber, etc.). Ultimately, semiconductor device makers must gauge the acceptable level of mask blank defects based on the built-in redundancy of their circuit designs. A few defects in the memory areas of circuits are generally acceptable, while even a single defect in the logic areas is usually unacceptable. Determining what defects will print is a complex exercise, and a complete discussion of this topic is beyond the scope of this text, but generally any defect larger than the half pitch of the densest printed circuit patterns is relevant and should be eliminated. Of concern are also defects that may cause a circuit to fail prematurely. For example, a defect that causes a conductor to be undersized may permit the circuit to work initially, but may lead to infant mortality. This is the traditional reason why redundancy alone in memory chips does not result in acceptable defects.

Substrate defects tend to fall into two categories: particles and pits (see Fig. 2.11). Both result from the polishing processes used to achieve atomic smoothness in the surface of the LTEM material. While cleaning can remove added particles (prior to multilayer deposition, of course), pit defects are harder to fix. Local polish touch-up and ‘smoothing’ during the multilayer deposition process can reduce the depth of pit defects, limiting their potential to print, but generally pit defects should be addressed through prevention. Also, a small pit may result in a large defect after the multilayer is deposited.

The multilayer deposition process results in several types of defects (see Fig. 2.12), which can be broadly classified as amplitude or phase defects.



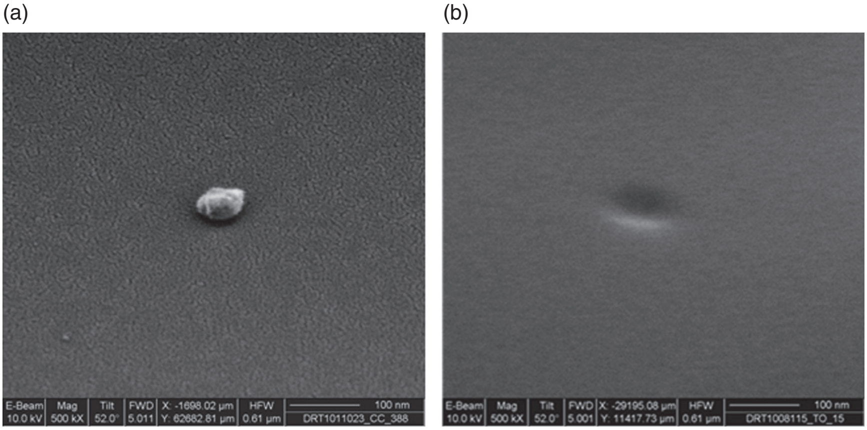
2.12 EUV mask blank defects.

An amplitude defect directly absorbs EUV light, creating an intensity difference, and generally results from disruptions (embedded particles, cracks, etc.) near the top of the multilayer. A phase defect results when reflections from successive multilayers fail to constructively interfere due to phase difference, and occurs when the multilayer pattern has swelled or dipped slightly at the surface. Both amplitude and phase defects create imaging artifacts when the mask is exposed in an EUV lithography tool.

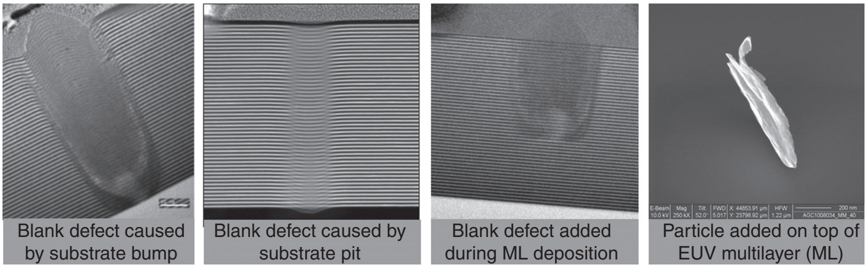
Amplitude and phase defects are detected using different techniques. Amplitude defects – surface or near-surface defects – can generally be detected with optical wavelength inspection tools. Phase defects – buried defects – may not be detected with optical inspection tools and generally must be detected using actinic, or at-wavelength, metrology. In the case of EUV mask blank phase defects, this requires the use of an EUV photon-based inspection tool. The inspection tools that find defects on the EUV mask blanks generally perform one mask inspection in one hour, and have detection capabilities down to about 30 nm (on the mask). The best defect levels for EUV mask blanks commercially achieved are on the order of 10–20 defects at a size of 50 nm or larger, with zero defects above 100 nm.

Defect repair

The repair of defects in the absorber layer is done using the same techniques employed to repair chrome-on-glass masks used in optical lithography. One added concern in the case of EUV stems from the non-normal incidence of the EUV light. Specifically, the height of the repaired feature


(a)
(b)
E-Beam 10.0 kV Mag 500 kX Tilt 52.0° FWD 5.011 X: -1698.02 µm Y: 62682.81 µm HFW 0.61 µm 100 nm DRT1011023_CC_388
E-Beam 10.0 kV Mag 500 kX Tilt 52.0° FWD 5.001 X: -29195.08 µm Y: 11417.73 µm HFW 0.61 µm 100 nm DRT1008115_TO_15

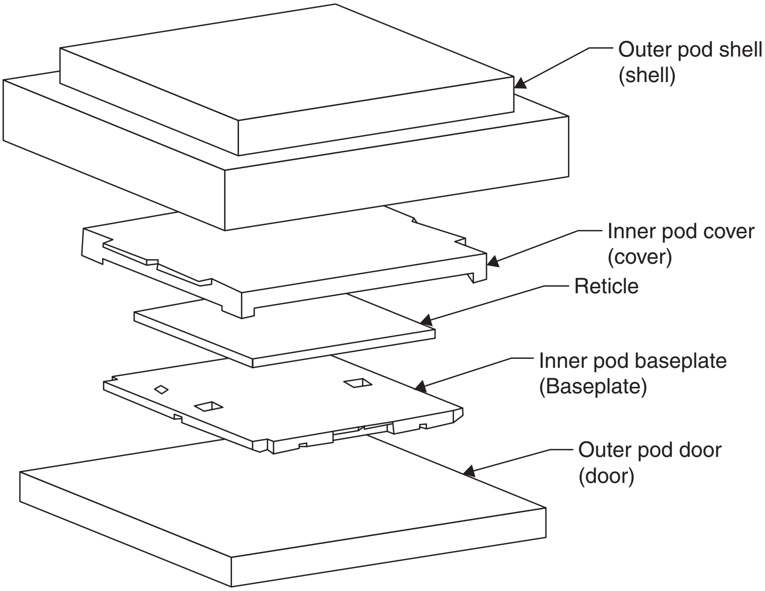
2.13 A repaired absorber layer defect of the proper height and one with excess height exhibiting a shadowing effect. (Image courtesy Tony Yen, TSMC.)


Blank defect caused by substrate bump
Blank defect caused by substrate pit
Blank defect added during ML deposition
Particle added on top of EUV multilayer (ML)

2.14 Opaque and clear defects and their repair on an EUV mask. (Image courtesy Tony Yen, TSMC.)

must be approximately correct in order to avoid creating a significant shadowing effect (see Fig. 2.13). Opaque (excess absorber material) and clear (missing absorber material) absorber defects can be repaired equally well (see Fig. 2.14). The repair process is quite lengthy, however, typically requiring about 1 h per defect. At today's defect levels, this equates to roughly one to two masks repaired per day per repair tool.

Defects in the multilayer cannot be reliably repaired as of writing this. Multilayer defects that are covered by absorber material, however, are effectively eliminated. One countermeasure for multilayer defects, then, is to determine if there exists a placement of the pattern relative to the blank that covers all multilayer defects with absorber. This strategy is generally effective only for a very small number of multilayer defects. The implication is that if a mask blank has very many multilayer defects, it will most probably be unusable.



2.15 An EUV Pod from the SEMI E152 specification. (Used with by permission from SEMI®.)

2.4.3 EUV mask handling

After EUV masks have been fabricated they must be transported to semiconductor manufacturing facilities, loaded and unloaded into exposure equipment hundreds or thousands of times, and exposed millions of times. At each step in the process the mask is susceptible to the addition of defects. The need to prevent particle adders led to the creation of a mask carrier specification, SEMI E152. The specification calls for an EUV Pod (see Fig. 2.15), which consists of two pods – an outer pod that is handled in 'dirty' environments (i.e. the reticle storage area, the cleanroom, etc.), and an inner pod that is handled in 'clean' environments (the ultra-clean environment inside the scanner or the mask cleaning tool). The inner pod behaves like a virtual pellicle, protecting the surface of the mask from particles during transport of the mask.

Chemical contamination in the form of carbonization or oxidation is also a potential hazard. For this reason, the SEMI E152 spec calls for the dual-pod to be purged during storage and transport to maintain an inert environment.

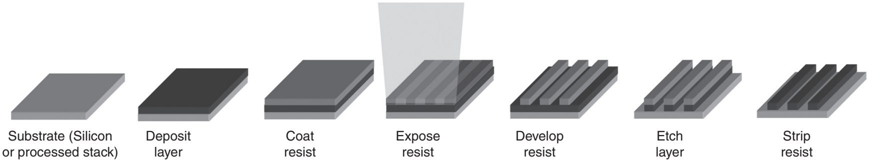
Inside the EUV imaging equipment the mask must be removed from the inner pod and placed on a reticle stage prior to exposure. Although the mask surface is now completely exposed, one mitigation that is employed is to face the mask down, exploiting gravity and the absence of air flow (the scanner environment is high vacuum) to eliminate those (larger) defects that would otherwise 'fall' onto the surface. Electrostatic forces, however, are much larger than gravity for particles a few nanometers in size.

2.5 EUV resists

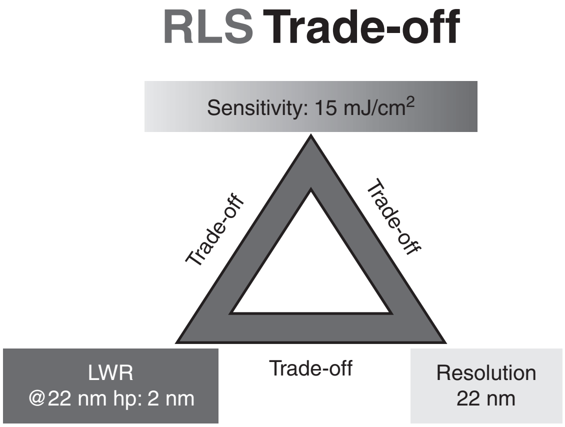
The object of lithography is to transfer circuit patterns from a mask to semiconductor devices in various states of fabrication. Figure 2.16 depicts the lithographic process. The material to be patterned is a thin film deposited onto a circular wafer with a silicon substrate. The wafer is coated with a photo-sensitive chemical emulsion called photoresist. The resist-coated wafer is exposed with light in the shape of a circuit pattern, chemically altering the illuminated resist. The figure depicts a positive-tone resist process where exposed photoresist is removed by a developer. Negative tone resist processes, which remove the unexposed material, can also be used. The patterned resist-coated wafer is then etched, removing the underlying material wherever the resist is absent. The lithographic transfer process is complete after a final chemical stripping process to remove the resist. Strictly speaking, the photolithography process comprises the middle three steps in the diagram: coat, expose, develop.

After commercial EUV efforts began, resist makers found that 248 nm resist systems exhibited reasonable lithographic performance under exposure to 13.5 nm light. These resist systems were chemically amplified: a polymer solution contains a photo-acid generator that catalyzes multiple deprotection reactions in a base polymer, making exposed regions of the film developer-soluble. 50 nm lines were resolved using chemically amplified resist on the SEMATECH LBNL MET in 2004 (effectively matching results achieved in 1990 on non-chemically amplified resist), but the initial insertion target for EUV at around 45 nm features was missed, due to the general lack of readiness of EUV technology, and it became necessary to develop higher resolution materials.

A detailed discussion of the chemistry of photoresist in general can be found elsewhere in this text. This section focuses on the requirements imposed on EUV resists, and the challenges specific to creating commercial, high resolution photoresists for EUV.



2.16 The photolithography pattern transfer process.



2.17 The resolution:LWR:sensitivity, or RLS, trade-off at the 22 nm node specification.

2.5.1 EUV resist requirements

Today, at a minimum, EUV resist performance must meet 22 nm resolution, 2 nm line-width roughness (LWR), and a sensitivity of 15 mJ/cm^2. (Long-term resolution must be improved to 8 nm or better and LWR to 1 nm.) These three requirements make competing demands upon the resist system. This is known as the RLS trade-off (Gallatin, 2005) and depicted in Fig. 2.17. Two of the three attributes can typically be improved at the expense of the third. In optical lithography, the stage-limited nature of the exposure systems means that a surfeit of photons is available to be used, relaxing the sensitivity requirements and allowing resolution and LWR to improve. For EUV resists, however, EUV exposure tools are currently photon-limited. The result is that EUV resists must be made to expose with a minimum of EUV photons, removing that particular design freedom.

2.5.2 EUV resist challenges

After the initial success of 248 nm resist systems in the early 2000s, the requirements on EUV resist rapidly exceeded their performance. Updated insertion timings for EUV technology suggested that 22 nm or better resolutions would be required with a roadmap to even better ultimate performance. Optimization of the resist systems revealed a number of significant challenges.

Because of their energy (92 eV), EUV photons interact with the atomic electrons in resist films rather than with the molecules. As a result, the absorbance of the material derives from the elemental formulae of the resist materials and their densities, rather than from the molecular structure of the constituents. The consequence of high absorbance is a non-uniform exposure of the film, potentially causing underexposure at the bottom of the resist (for positive resists) and resulting in angled resist profiles (wider at the bottom than at the top). Too low an absorbance can lead to a particularly poor sensitivity and, in the case of positive resists, a re-entrant profile (wider at the top than at the bottom).

The need to resolve extremely small features also places restrictions on the thickness of resist that can be employed. The so-called aspect ratio (AR) is defined as the height divided by the width of the features being patterned. Historically, ARs of 2.5 or less have been acceptable. Higher ARs begin to incur larger potential for pattern collapse. Film thicknesses around 40 nm are routinely used when resolving ~20 nm features today.

Since EUV photons are highly energetic and scarce, it is necessary to make the most efficient use of them possible. In an optical system, a given photon may interact with a photo-acid generator (PAG) and generate a single acid, or it may not interact at all. This rate of interaction is called the quantum yield, and for optical systems it is rigorously capped at 1. EUV photons, with 92 eV, tend to distribute their energy to a number of electrons, each of which has a chance to interact with the PAG in the film. The quantum yield of EUV systems can therefore exceed 1.

This potential for high quantum yields is welcome news for EUV resist, as it tends to counterbalance a statistical phenomenon called shot noise. Shot noise refers to the statistical noise present when imaging a feature. This variation of delivered exposure dose to a given volume scales as $N^{1/2}$, where N is the number of photons being delivered to a given volume (Brainard *et al.*, 2004). The shot noise is proportional to $1/N^{1/2}$. If the number of contributing photons is small, the variation is correspondingly large. Using fewer photons to expose material will result in significant shot noise and a rough-looking resist feature.

As mentioned in Section 2.3.4, hydrocarbons can contaminate EUV optical surfaces, reducing their reflectivity. Unfortunately, EUV photons are energetic enough to crack molecules present in resists, resulting in the evolution of hydrocarbon-containing products in a process called resist outgassing. Strict limits are placed on the amount of outgassing permitted from photoresists. Materials that exceed this requirement cannot be used in commercial exposure systems.

Finally, chemically amplified EUV resists depend upon a diffusion-based acid deprotection process. For each acid generated there is a finite spread of molecular reactions. This ‘blurring’ effect tends to increase the roughness of the feature. The standard prescription for improving LWR is to reduce the acid diffusion length, effectively reducing the number of deprotection reactions catalyzed by each acid molecule. This reduces the ‘amplification’ of the chemically amplified resist, making it less sensitive and requiring more light to expose. Since EUV is photon-limited, this approach is problematic.

2.5.3 EUV resist approaches and progress

The complex set of requirements upon EUV resists have resulted in the exploration of a number of innovative approaches. Many variations of chemically amplified photoresists have been explored for use in EUV. Polyhydroxystyrene (PHS) resists were the most common of these materials used in the early and mid 2000s. As the resolution requirement tightened below 30 nm, researchers began experimenting with radical alternatives: non-chemically amplified resists, chain scission resists, molecular glasses, and many others. In some cases, additional elements were added to the resist system, such as photo-base-generator and polymer-bound PAGs.

EUV resist development is highly competitive. Many recent enhancements being added to photoresists are proprietary, and are not publicly disclosed. Progress over the past decade is shown in Fig. 2.3, with 16 nm (and even better) resolution achieved using EUV tools with a 0.3 NA. The key remaining challenge for EUV resists remains the LWR, which generally exceeds 3 nm (3σ) vs a requirement of less than 2 nm (3σ). Ultimately, the LWR of consequence is that of the pattern transferred to the substrate. For this reason, additional LWR improvement may come from resist post-processing techniques that smooth the resist features after they are developed.

2.6 EUV integration and implementation challenges

Producing an EUV exposure system that can be used in a manufacturing environment is a formidable task. First, it requires that solutions be available for all of the technological challenges introduced in the preceding sections. But additionally, these solutions must be integrated with each other seamlessly in a highly reliable, cost-effective product. A few of these integration challenges are highlighted in this section.

2.6.1 Out-of-band radiation

Light from the EUV source is directed into an intermediate focus. The light should be a 2% bandwidth of EUV photons centered on 13.5 nm, but because of the plasma nature of the high power source technology, the spectrum found at intermediate focus is much broader. This OOB radiation can act as a contrast-reducing background light if it is transmitted through the optical train to the photoresist, leading to degradation in performance.

Fortunately, it is not necessary to block all OOB radiation, since not all light will expose the photoresist, and the optical system acts like a band pass filter already. The wavelengths that must be blocked are restricted to ultraviolet (UV) light in the 130 nm to 400 nm range, and visible light > 400 nm. The ultimate goal is to reduce OOB to a fraction (3–7%) of the in-band power.

System integrators use a spectral purity filter (SPF) to achieve these goals. A variety of SPF approaches exist. One common approach is to use a thin film that is highly reflective in the UV and IR. Any thin film will result in loss of EUV light and heating of the film (which could destroy it), so minimizing the thickness is of paramount importance. However, if the film is too thin it will lack structural strength. An alternative approach is to coat one or more of the optics in the illuminator with films that minimize transmission of unwanted spectral bands through destructive interference. Still another involves the use of a blazed grating that reflects UV and IR light at a different angle than the EUV light.

2.6.2 System contamination

As mentioned in Sections 2.3.4 and 2.4.3, carbon deposition can degrade the reflectivity of EUV optics. The countermeasure is to carefully restrict carbon emitters from being present in areas that EUV light may reach. This is simple in principle, but extremely challenging to achieve in practice. EUV system integrators found that even when system components were fabricated from approved materials, the fabrication processes themselves often used proscribed materials to machine, polish, clean, etc., the components, thereby contaminating them!

Components and sub-components used in EUV equipment must therefore be scrutinized and reviewed to ensure strict compliance from start to finish. For example, it is not uncommon for component manufacturers to be required to purchase duplicate equipment for use only in the manufacture of parts destined for EUV tooling.

Despite the best efforts at prevention, some background pressure of hydrocarbons is inevitable. Combined with the long lifetime mandated for the optics, an *in situ* cleaning method is a necessity. As mentioned in Section 2.3.4, reactive gas at low pressure can remove deposited carbon contamination.

2.6.3 System weight and power

Fully integrated EUV exposure systems are some of the most massive pieces equipment ever used in semiconductor manufacturing, approaching 50 tons! The vacuum system and stages are the primary contributors to the system weight. The vacuum vessel, which contains the entire optical train and mask and wafers stages, must be stiff enough to remain rigid to very precise tolerances despite having an extremely large surface area. The stages move at speeds of 0.5 m/s or faster, but must be able to place images to nanometer accuracy, a feat that requires a significant mass to provide stability.

The floors of most modern semiconductor fabrication facilities (fabs) were not designed to support weights as high as those of EUV tools. New fabs are being designed to increase the tolerance of the floors in the lithography sections, while older fabs must be retrofitted to add support to handle the weight.

Power consumption from EUV tools is also significant. LPP sources employ banks of high power lasers. *In situ* optics cleans involving hydrogen gas require abatement systems that can draw significant amounts of power as well. A production EUV exposure system is projected to consume 150 kW or more.

2.7 Conclusion and future trends

EUV lithography is an extremely complex technology, with a number of difficult challenges still to be overcome before it can be successfully commercialized. Significant resources will be required to see EUV broadly implemented in high volume manufacturing .

2.7.1 Benefits of EUV

As mentioned in the introduction, the theoretical limit of single exposure patterning using 193 nm immersion scanners is about 36 nm. It is, however, possible to further reduce this with the use of additional processing steps that alter the feature size (although not the pitch) and interleave subsequent optical patterns in a process known as double patterning. As of this writing, 22 nm features are being fabricated in commercial environments using just such approaches. It is further possible to continue this approach beyond ‘double’ patterning into ‘multiple’ patterning at the cost of additional processing complexities and lithography steps.

Despite the evident promise of double patterning and multiple patterning, the semiconductor industry is forced to recognize that doubling (or tripling, or quadrupling, etc.) the number of processing steps also doubles (or triples, or quadruples, etc.) the cost of the solution. The addition of each step introduces more potential for yield loss as well, exacerbating the cost impact of the added processing steps. For this reason the industry prefers a single exposure solution.

In the 1990s, many lithography technologies were considered as alternatives to optical lithography. Proximity X-ray, ion projection lithography, e-beam lithography, and soft X-ray projection lithography were some of the approaches studied and even commercialized across the globe. The saga of the choice of a so-called 'next-generation lithography' is beyond the scope of this text. Although EUV lithography's success as a manufacturing solution has not been determined, the semiconductor industry has invested more broadly in that than in any other next-generation lithography. The reasons for this industry decision bear deeper inspection.

The clear benefit of EUV is a 15× reduction in the wavelength, compared to today's state-of-the-art 193 nm optical lithography equipment. Since Moore's law is couched in terms of 'doubling' transistor densities every two years (a so-called generation or 'node'), such advancement in resolution would provide several generations of technology extendibility.

Ultimately, the choice of EUV lithography to succeed optical lithography is being made for economic reasons: the technology promises several generations of usefulness during which the cost-effective 'single exposure' approach can be continued. Semiconductor device makers can use the attendant increase in functionality resulting from continued transistor scaling to sell better, more powerful products for equal or better revenue. This business model breaks without the clockwork advancement of technology. Rather than face such an unattractive possibility, the industry is willing to make heavy investments.

2.7.2 Limits of EUV

EUV lithography exposure tools with NA = 0.5 will be able to pattern 8 nm features in a single exposure. The next process node, at 5.5 nm feature sizes, would require either a wavelength shift or the implementation of EUV double patterning. If a wavelength shift approach is adopted by the industry, 6.8 nm has been proposed.

It is not clear when the semiconductor industry will reach the end of Moore's law. At some point, traditional lithographic techniques will have to make way for radical new technologies. Atomic scale transistors, self-aligned chemistries, and other radical approaches will undoubtedly emerge, challenging the semiconductor industry to adapt and evolve in order to succeed. But for now EUV, if successfully implemented, promises to enable the continued extension of semiconductor scaling through to the end of the 2010s.

2.8 Acknowledgments

The author would like to thank the following individuals for valuable discussions that made this chapter possible: Dominic Ashworth, Terence Bacuita, Robert Brainard, Robert Bristol, Manish Chandhok, Greg Denbeaux, Kent Frasure, Michael Goldstein, Frank Goodwin, Erik Gullikson, Ahmed Hassanein, Long He, Vibhu Jindal, Patrick Kearney, Chris Krautschik, Andy Ma, Hans Meiling, Warren Montgomery, Patrick Naulleau, Eric Panning, Abbas Rastegar, Jeanette Roberts, Gil Vandentop, Stefan Wurm, Tony Yen, and Paul Zimmerman.

Finally, this work could not have been completed without the support of my wife, Katie, who endured many nights and weekends of writing and remained good natured throughout.

2.9 References

Bajt S, Alameda J B, Barbee T W Jr., Clift W M, Folta J A, Kaufmann B, Spiller E A (2002), 'Improved reflectance and stability of Mo-Si multilayers', *Opt. Eng.*, **41**, 1797.

Berreman D W (1991), 'Multilayer reflecting x-ray optical systems: chromatic vignetting by narrow reflection bands', *Appl. Opt.*, **30**, 1741.

Berreman D W, Bjorkholm J, Eichner L, Freeman R R, Jewell T E, Mansfield W M, MacDowell A A, O'Malley M L, Raab E L, Silfvast W T, Szeto L H, Tennant D M, Waskiewicz W K, White D L, Windt D L, Wood II O R and Bruning J H (1990), 'Soft x-ray projection lithography printing 0.20 mm features using 20:1 reduction', *Opt. Lett.*, **15**, 529–531.

Bijkerk F, Voorma H-J, Puik E J, Louis E, van Dorssen G E, van der Wiel M J, Verhoeven J, van der Drift E W J M, Romijn J, Rousseeuw B A C (1991), 'Design of an extended image field soft-x-ray projection system', in *OSA Proc. on Soft-X-Ray Projection Lithography*, **12**, 51–53.

Bjorkholm J, Bokor J, Eichner L, Freeman R R, Gregus J, Jewell T E, Mansfield W M, MacDowell A A, Raab E L, Silfvast W T (1990), 'Reduction imaging using multilayer coated optics: printing of features smaller than 0.1 microns', *J. Vac. Sci. Technol. B*, **8**, 1509–1513.

Brainard R L, Trefonas P M, Lammers J H, Cutler C, Mackevich J, Trefonas A and Robertson S (2004), 'Shot noise, LER, and quantum efficiency of EUV photoresists', *Proc. SPIE*, **5374**, 74–85.

Chen A (2010), 'EUV exposure systems for volume production', 3rd Annual SEMATECH Symposium Taiwan 2010, Hsinchu, Taiwan, 3–9 September 2010.

Corthout M and Neff W (2007), 'The philips extreme UV Sn source: on the way to an integrated system for high volume manufacturing', EUV Source Workshop, Baltimore, US.

Feldmann H, Ruoff J, Harnisch W and Kaiser W (2010), 'Actinic review of EUV masks', *Proc. SPIE*, **7636**, 76361C.

Fujimoto J, Hori T, Yanagida T, Ohta T, Kawasuji Y, Shiraishi Y, Abe T, Kodama T, Nakarai H, Yamazaki T and Mizoguchi H (2012), 'Development of laser-produced plasma-based EUV light source technology for HVM EUV lithography', *Proc. SPIE*, **8322**, 83220F.

Gallatin G M (2005), 'Resist blur and line edge roughness', *Proc. SPIE*, **5754**, 38.

Hassanein A, Sizyuk V, Morozov V, *et al.*(2005), 'Dynamic simulation of discharge and laser produced plasma for EUV lithography devices', IEEE Conference

Record – Abstracts. 2005 IEEE International Conference on Plasma Science – ICOPS 2005 (IEEE Cat No. 05CH37707): 103.

Hawryluk A M and Seppala L G (1998), 'Soft x-ray projection lithography using an x-ray reduction camera', *J. Vac. Sci. Technol B*, **6**, 2162.

Hawryluk A, Ceglio N and Gaines D (1989), 'Reflection mask technologies for x-ray projection lithography', *J. Vac. Sci. Technol. B*, **7**, 1702.

Hawryluk A M and Ceglio N M (1993), 'Wavelength considerations in soft-x-ray projection lithography', *Appl. Opt.*, **32**, 7062.

Henke B L, Gullikson E M and Davis J C (1993), 'X-ray interactions: photoabsorption, scattering, transmission, and reflection at E = 50–30 000 eV, Z = 1–92', *At. Data Nucl. Data Tables*, **54**, 181.

Jurczyk B E, Alman D A, Antonsen E L, Jaworski M A, Williams M J, Ruzic D N, Spila T, Edwards G, Wurm S, Wood O, Bristol R L (2005), 'The effect of debris on collector optics, its mitigation and repair: next-step a gaseous Sn EUV DPP source', *Proc. SPIE*, **5751**, 572.

Kinoshita H, Kurihara K, Ishii Y and Torii Y (1989), 'Soft x-ray reduction lithography using multilayer mirrors', *J. Vac. Sci. Technol. B*, **7**, 1648.

Kirkpatrick P and Baez A V (1948), 'Formation of optical images by X-rays', J. *Opt. Soc. Am.*, **38**, 766–773.

Kortright J B, Gullikson E M and Denham P E (1993), 'Masked deposition techniques for achieving multilater period variations required for short-wavelength (68-A) soft-x-ray imaging optic', *Appl. Opt.*, **32**, 6961.

Kubiak G D, Tichenor D A, Malinowski M E, *et al.* (1991), 'Diffraction-limited soft x-ray projection lithography with a laser plasma source', *J. Vac. Sci. Technol. B*, **9**, 3184.

Kuerz P, Boehm T, Dinger U, Mann H, Muellender S, Dahl M, Lowisch M, Muehlbeyer M, Natt O, Rennon S, Sohmen E, Stein T, Wittich G, Kaiser W and Rupp W (2008), 'Optics for EUV Lithography', 2008 International EUVL Symposium, Lake Tahoe, CA, 28 September-1 October 2008.

Lee S, Byers J, Jen K, Zimmerman P, Rice B J, Turro N J and Willson C G (2008), 'An analysis of double exposure lithography options', *Proc. SPIE*, **6924**, 69242A.

Litt L C, Bourov A, LaFontaine B and Apelgren E (2002), 'Evaluation and characterization of flare in ArF lithography', *Proc. SPIE*, **4691**, 1442.

Loewen R J (2003), 'A Compact Light Source: Design and Technical Feasibility Study of a Laser-Electron Storage Ring X-Ray Source', Ph.D. Dissertation, Stanford University, SLAC-R-632.

Lowisch M, Kuerz P, Boehm T, *et al.* (2007), 'Optics for EUV Lithography', 2007 International EUVL Symposium, Sapporo, Japan, 28–31 October 2007.

Rettig C L, Khodykin O V, Hoffman J R, Marx W F, Bowering N R, Vargas E, Ershov A I, Fomenkov I V and Partlo W N (2005), 'Protection of collector optics in an LPP based EUV source', *Proc. SPIE*, **5751**, 910.

Richardson M, Torres D, DePriest C, Jin F and Shimkaveg G (1998), 'Mass-limited, debris-free laser-plasma EUV source', *Op. Comm*, **145**, 109.

Silfvast W T and Wood O R II (1988), 'Tenth micron lithography with a 10 Hz 37.2 nm sodium laser', *Microelectron. Eng.*, **8**, 3–11.

Skulina K, Alford C, Bionta R, Makowiecki D, Gullikson E M, Soufli R, Kortright J B and Underwood J H (1995), 'Beryllium based multilayers for normal incidence EUV reflectance', *OSA Proc. on Extreme Ultraviolet Lithography 1994*, **23**, 52–55.

Spiller E (1994), *Soft X-Ray Optics*, SPIE Press, Bellingham, WA.

Stearns D G, Rosen R S, and Vernon S P (1993), 'Multilayer mirror technology for soft-x-ray projection lithography', *Appl. Opt.*, **32**, 6952.

Uda K, Mizusawa N, Tanaka Y, Watanabe Y, Ina H, and Uzawa S (1998), 'X-ray stepper development for volume production at Canon', *Proc. SPIE*, **3331**, 689.

Underwood J H and Barbee T W (1981), 'Layered synthetic microstructures as Bragg diffractors for X rays and extreme ultraviolet: theory and predicted performance', *Appl. Opt.*, **20**, 3027.

Wood O R II, Bjorkholm J E, Dreyer K F, *et al.* (1995), 'Experiments and simulations of EUV lithographic resist patterning at wavelengths from 7 to 40 nm', *OSA Proc. on Extreme Ultraviolet Lithography*, **23**, 83–88.

Xie P and Smith B (2009), 'Analysis of higher order pitch division for sub-32 nm lithography', *Proc. SPIE* **7274**, 72741Y.

Zernike F and D. Attwood (1994), 'Extreme ultraviolet lithography', Eds., *OSA Proc.*, **23**.

Zimmerman P A, Rice B J, Piscani E C and Liberman V (2009), 'High index 193 nm immersion lithography: the beginning or the end of the road', *Proc. SPIE*, **7274**, 727420.

3

Electron beam lithography

T. R. GROVES, University at Albany (SUNY), USA

DOI: 10.1533/9780857098757.80

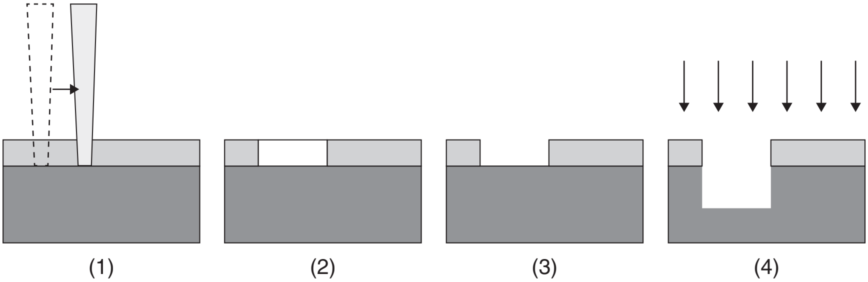
Abstract: A focused electron beam (e-beam) represents the smallest, finest practical writing pencil known, with the capability of producing pattern features down to a few nanometers in size. Electron beam lithography does not rely on a pre-existing patterned mask, but can write the pattern directly from stored data. Because of its inherent high resolution and pattern flexibility, e-beam lithography remains the method of choice for fabricating nanometer-scale structures in low volume. The historical Achilles heel of e-beam lithography has been its low throughput. This can be mitigated by exposing many pixels in parallel. A survey of present-day e-beam lithography is presented.

Key words: electron beam, lithography, maskless, throughput, multiple beams.

3.1 Introduction

Electron beam lithography, also known as e-beam lithography, is the process of tracing out a pattern in a suitable recording medium using a focused e-beam. The underlying physical mechanism relies on the fact that the recording medium, typically a thin organic polymer film, is altered by the passage of fast electrons. The recording medium is generally called *resist*. In a subsequent development step, the exposed material is removed (positive-tone process). Alternatively, the unexposed material is removed (negative-tone process). In either case, the result is a patterned film, which acts as a binary mask for further processing. This processing might include reactive ion etching, selective ion implantation, electroplating, or physical vapor deposition, to name a few. The patterned binary mask is a versatile and inexpensive enabler for a variety of subsequent processes. By superimposing multiple pattern layers, an enormous variety of useful devices can be fabricated. A typical positive-tone process is shown schematically in Fig. 3.1.

A focused e-beam represents the smallest, finest practical writing pencil known (Pease and Chou, 2008; Pease, 2010). The ultimate electron optical resolution is the same as an electron microscope, in the range of 0.06–0.15 nm,



3.1 Typical positive-tone resist process. From left to right: electron beam scans (1) and leaves behind a latent image in the resist layer (2). Exposed resist (3) dissolves in the development process. Subsequent reactive ion etch (RIE) selectively removes substrate material (4), using the developed resist layer as a binary mask.

depending on the energy of the incident electrons. Ultimate lithographic resolution is not limited by the electron optics, but by the range of interaction of the beam electrons with the resist layer by scattering and secondary processes. Ultimate lithographic resolution is typically in the range of a few nanometers, depending on the energy of the electrons and the specific nature of the resist. This is about an order of magnitude smaller than the lithographic resolution obtainable with conventional optical lithography.

The pattern data are typically created using commercially available software for computer-aided design. These data must then be converted to a format usable by the e-beam writer. A digital electronic data path automatically converts and sends the data to the e-beam writer. The e-beam is then scanned over the writing surface using electric or magnetic fields, and turned on and off while it scans. Practically any arbitrary binary pattern can be written in this way.

In conventional optical lithography, one forms a demagnified image of a pre-existing patterned mask onto a resist-coated wafer. The mask can be used repeatedly to make many copies of the same pattern. In high-volume manufacturing of semiconductor chips, the patterns are highly complex. For example, 30 nm minimum-sized features might be distributed over a 30 mm square area. This represents an upper limit of 10^{12} pattern features. Consequently, the mask can be expensive and time-consuming to fabricate. It is only cost-effective if many wafers are exposed with a single mask, since the cost of the mask is amortized over all of the wafers exposed.

E-beam lithography does not require a pre-existing mask, since the pattern is created and transmitted electronically. This permits great flexibility in trying out a large number of different patterns in a short time. This is ideal for low-volume applications, in which few copies of a given pattern are needed. An e-beam writer is a pattern *generator*, whereas a conventional optical lithography tool is a pattern *replicator*. Incidentally, the method of choice for patterning masks for optical lithography is e-beam lithography.

An e-beam lithography system is comprised of several subsystems, including

- an electron optical column, to produce the focused e-beam;
- analog electronics to produce, focus, blank/unblank, and scan the beam;
- digital electronics to store and transmit the pattern data;
- a high-precision mechanical XY stage to position the writing substrate relative to the e-beam;
- a high-vacuum system, with provision to move the writing substrate in and out of the vacuum;
- high-speed computers and microprocessors, to automatically perform all of the necessary tasks;
- an extensive software system.

A considerable engineering effort is needed to make all of these components work reliably together.

E-beam lithography originally grew out of scanning electron microscopy. It was first proposed by Buck and Shoulders, (1958), and first demonstrated by Möllenstedt and Speidel (1960). An electron microscope *captures* a high-resolution image of a pre-existing object. An e-beam writer *creates* a structure with moderate spatial resolution (relative to an electron microscope), but high accuracy and an enormous number of pixels. A high-resolution electron micrograph typically has 10^6 pixels. A high-fidelity e-beam-written pattern can have 10^{12} pixels, with each pixel precisely positioned to within a reasonably small fraction of the pixel size. Consequently, an e-beam writer is much more complicated and expensive than an electron microscope.

The overriding goal in any lithographic patterning is to produce a pattern in resist that approximates the original design pattern with the greatest possible fidelity. Lithographic patterns as exposed in resist are binary. As such, they have no gray scales. (The exposure process can utilize gray scales, but the resist image is typically binary). In this context, pattern fidelity consists of two basic attributes. One is the quality of individual pattern features, as embodied in dimensional control of the feature size. This includes the smoothness of feature edges, and the sharpness of corners. The other is accurate placement of pattern features.

The written and processed pattern generally does not perfectly match the ideal, desired pattern. Errors in pattern feature size and placement can arise from multiple sources. For example, unwanted fluctuation of the exposure dose leads to non-uniformity in the printed feature size. Electromagnetic noise and mechanical jitter in the system lead to random errors in pattern placement.

Charging of contamination layers in the electron column, together with uncontrolled thermal expansion due to local temperature fluctuations, causes the beam to wander. Such errors are often not predictable or repeatable, and must be minimized by an iterative process of measurement, deduction (to determine the source of the error), and reduction to a tolerable level.

Some errors are predictable and repeatable, such as average deflection distortion, average position error of the mechanical stage, and average scattered dose variation arising from local variation in pattern density (so-called proximity effect). This class of errors is amenable to correction by automated measurement, computation, and feedback. Pattern fidelity is deemed acceptable if the errors in printed feature size and placement are a reasonably small fraction of the minimum printed feature size, also known as the *critical dimension* (CD).

In order to measure these errors, sophisticated metrology is needed as an indispensable adjunct to the writing process. Much of this metrology capability is built into the e-beam writer, and automated within the writing process. For example, the position of the e-beam relative to the writing surface can be measured by using the beam to scan alignment marks placed on the writing surface in a prior processing step. The position of the beam can thus be corrected to compensate for distortion and drift. In addition, a laser interferometer can be used to measure the position of the XY stage relative to a stable mechanical datum built into the system. The laser interferometer forms the built-in reference standard for all measurements pertaining to pattern placement. A typical laser interferometer for e-beam lithography has a resolution of $\lambda/1000$, where λ is the wavelength (632.8 nm) of a He-Ne laser. Higher resolution is also available. In principle, this resolution represents the smallest possible increment of pattern edge placement.

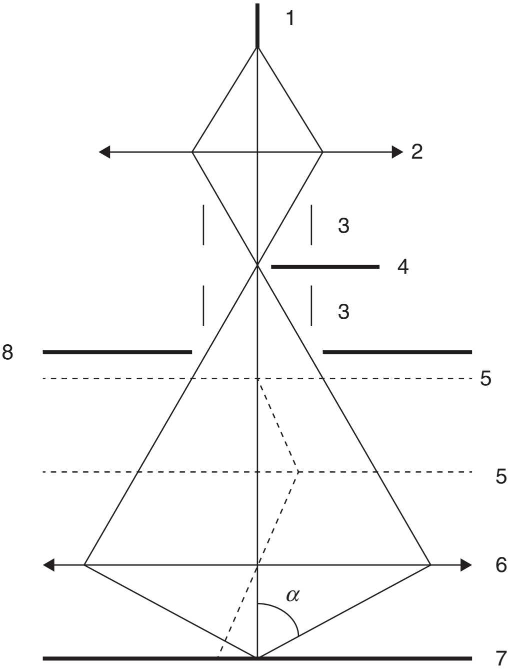
The historical Achilles heel of e-beam lithography is its slow speed. This arises from two limitations. First, the writing process is essentially serial, with the pattern traced out sequentially using a probe beam. This is in contrast to conventional optical lithography, in which an entire complex pattern is exposed in one flash or scan. Second, the useful writing current is limited. The beam electrons are randomly scattered by one another in the drift length of the electron column. This degrades the resolution as the current is increased. In addition, every electron optical system is limited in the amount of current it can supply at the writing surface for a given resolution. Useful writing current at any given resolution is always limited, either by Coulomb scattering in the beam path, or by limited ability of the electron optical system (especially the electron source) to supply the desired writing current.

A useful estimate of writing speed is the pattern area swept out per unit time by the beam. Typical e-beam writers operate in the range 0.0001–1.0 cm^2/s, depending on the desired resolution. By comparison, conventional optical exposure tools operate in the range 20–30 cm^2/s. Conventional optical systems are thus several orders of magnitude faster than currently existing e-beam systems. For this reason, optical lithography remains the method of choice for high-volume manufacturing of integrated circuit chips, while e-beam lithography remains the method of choice for device fabrication at low volume, with superior lithographic resolution.

In summary, e-beam lithography has the dual advantages of high spatial resolution and flexibility of pattern generation. It has the drawback of low speed. The purpose of this chapter is to explore these factors in some detail, and to offer some analysis of the possible avenues of future improvement.

3.2 Using pixel parallelism to address the throughput bottleneck

As mentioned, e-beam lithography systems use electric and magnetic electron lenses to form a sharply focused e-beam, which is scanned over the writing surface. The simplest possible configuration is one in which an image of a point-like electron source is formed directly on the writing surface (Chang *et al.*, 1976; Herriott *et al.,* 1975; Kelly *et al.,* 1981; Alles *et al.,* 1987).

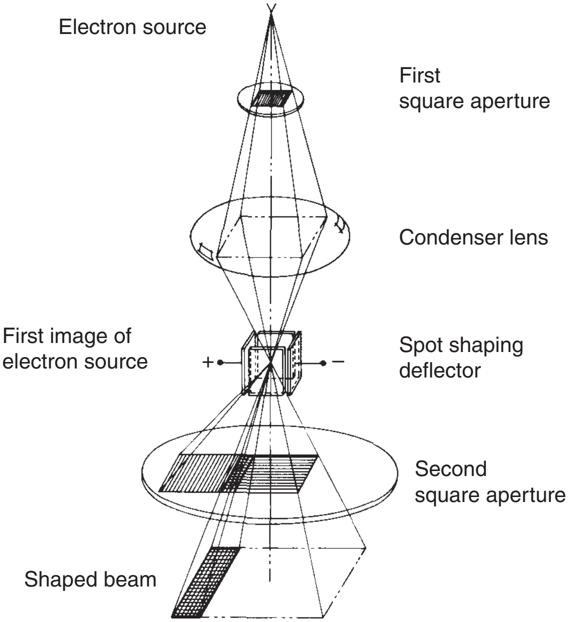


3.2 Typical Gaussian beam system configuration.

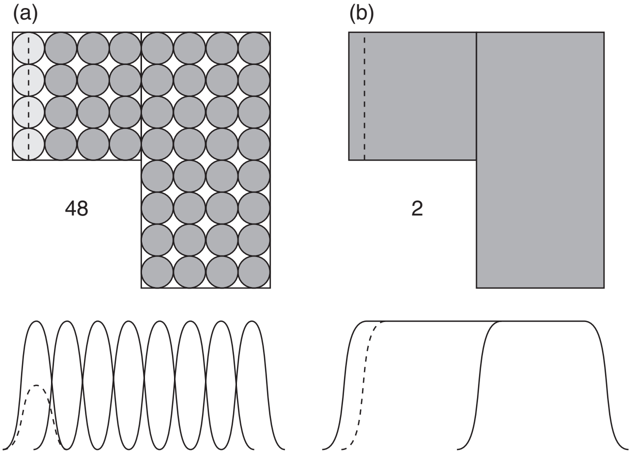
An example of this is shown schematically in Fig. 3.2. The compact electron source (1) is typically a thermally assisted field-emitter, also known as a Schottky emitter. The virtual source, seen 'looking back' toward the source from the electron column, typically has a diameter in the range 15–20 nm. A magnetic lens (2) forms an intermediate image of the source (1) in a set of electrostatic deflection plates (3). The deflection plates (3) move the beam onto an edge (4) for blanking. A magnetic lens (6) forms a demagnified image of the source onto the writing surface (7). Electromagnetic or electrostatic deflectors (5) move the writing spot laterally on the writing surface (7). The writing surface (7) is mounted on a movable stage for increased range of motion. An aperture (8) limits the illumination cone semi-angle α of the beam measured at the writing surface. An optimum aperture size exists for a given resolution that balances the effects of spherical aberration, diffraction, and Coulomb scattering within the beam. These effects will be described in more detail later.

The writing spot represents one pixel, and has a lateral intensity distribution that is roughly Gaussian. It is scanned over the pattern area using variable electric or magnetic fields, and turned off and on as it is scanned, thus generating the pattern. The beam can be scanned in a *raster* pattern, similar to a scanning electron microscope. Alternatively, the beam can be deflected only to those places where pattern features are to be written. This latter approach goes by the term *vector* scanning. Writing one pixel at a time is the ultimate serial writing process. Its inherent simplicity comes with the penalty of limited speed.

An alternative writing strategy is to project a rectangular writing spot of variable size and aspect ratio in a single flash (Pfeiffer and Loeffler, 1970; Pfeiffer, 1978; Goto *et al.*, 1978; Trotel, 1978). It is shown schematically in Fig. 3.3. The beam from an electron source floods a square aperture. The source is typically an extended (as opposed to point-like) single-crystal lanthanum hexaboride LaB_6 thermionic emitter. The extended source insures that the first square aperture is illuminated uniformly. A condenser lens forms an electron optical image of the square aperture in the plane of a second square aperture. A spot-shaping deflector moves the image of the first square aperture on the second square aperture, thus forming a rectangular compound spot of variable size and aspect ratio. The spot-shaping deflector is positioned at an intermediate image of the source (1). This insures that the illumination remains uniform as the beam is deflected. The resulting rectangular spot is then demagnified, and imaged onto the writing surface. In this way, many pixels can be exposed in a single flash, thus increasing the pixel parallelism. This is called a *variable-shaped beam* approach. Pixel parallelism varies with the size of the shape, but typically 64–256 pixels are written in a single 'flash.' Obviously, this is a significant increase over the single-pixel Gaussian beam.


Electron source
First square aperture
Condenser lens
First image of electron source
Spot shaping deflector
+
−
Second square aperture
Shaped beam

3.3 Typical variable-shaped beam system configuration.



3.4 Exposure flashes for a Gaussian beam system (a), and a variable-shaped beam system (b). The intensity distribution is plotted as a function of lateral position below. Incremental variation in pattern edge placement is also indicated.

It is possible to place the edge of a pattern feature with a precision that is a small fraction of the pixel resolution. This is shown schematically in Fig. 3.4. The pattern of writing for individual flashes is shown for a Gaussian beam on

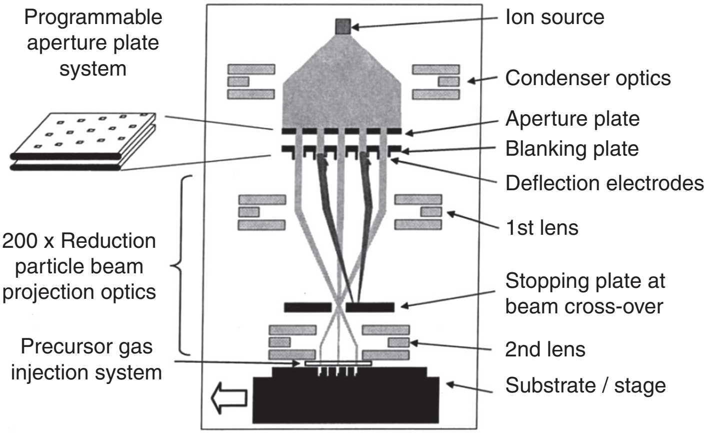
Fig. 3.4a, and for a variable-shaped beam on Fig. 3.4b. The number of individual flashes is indicated in both cases. We assume that the pixel resolution is the same for both cases. We now desire to move the leftmost pattern edge by an increment that is a small fraction of the pixel resolution. With a Gaussian beam, this can be done by exposing the leftmost column of pixels with reduced intensity. This will cause the edge of the resist image to move incrementally to the right, depending on the amount of exposure. This is referred to as *gray-scale* writing. This can also be used in conjunction with deflecting the beam by a small increment (Abboud 1997). With a variable-shaped beam, the placement of pattern edges is determined by the shaping increment, which is decoupled from the pixel resolution. This ability to place pattern edges with an increment that is a small fraction of the pixel size adds complexity to the writing, but it enhances the ability to write an enormous variety of useful patterns.

One can add further pixel parallelism by projecting a character or cell in a single flash. This is similar to the variable-shaped beam approach, with character apertures replacing the square apertures. The character is automatically selected from a library of shapes that reside in the electron column in the form of small stencil masks. Each stencil can be electron-optically imaged at will onto the writing surface. This approach, called *cell* or *character* projection (Pfeiffer 1979), is especially useful in applications for which a small pattern is repeated many times. In practice, this roughly doubles the pixel parallelism of the variable-shaped beam approach.

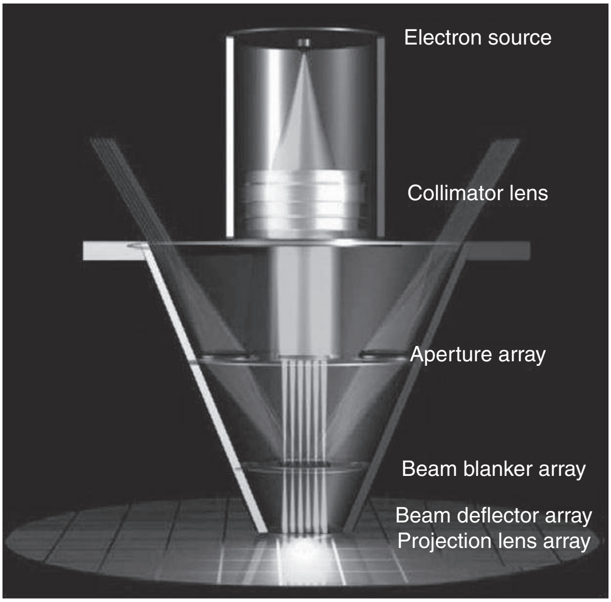
One can add still further pixel parallelism by using multiple e-beams in a single electron column. A single electron source floods an array of apertures, forming an array of individual beamlets. Below each aperture is a pair of electrostatic deflection plates, which is used to steer the beamlet onto a downstream aperture, thus blanking only that individual beamlet. It is necessary to address each pair of deflection plates individually.

In Fig. 3.5 (Platzgummer *et al.*, 2008) a single source of ions (or electrons) is collimated by the condenser optics. The beam floods an aperture array. Individual beamlets are either transmitted or blanked, consistent with the pattern pixels. The entire aperture array is imaged and demagnified by a system of lenses onto the writing surface below. The spacing between beamlets is demagnified by the same factor as the beamlet spot size. All of the beamlets pass through two intermediate crossovers in the column. An array of Gaussian beamlets is formed at the writing surface, each of which can be turned on and off at will. This approach is referred to as *projection maskless lithography patterning* (PMLP).

In Fig. 3.6 (Kruit, 1998; Wieland *et al.*, 2001; van den Berg *et al.*, 2011) each beamlet is individually imaged and demagnified onto the writing surface below. The spacing between beamlets does not change, and the individual beamlets do not pass through common crossovers. An array of Gaussian beamlets is formed at the writing surface, each of which can be turned on


Programmable aperture plate system
200 x Reduction particle beam projection optics
Precursor gas injection system
Ion source
Condenser optics
Aperture plate
Blanking plate
Deflection electrodes
1st lens
Stopping plate at beam cross-over
2nd lens
Substrate / stage

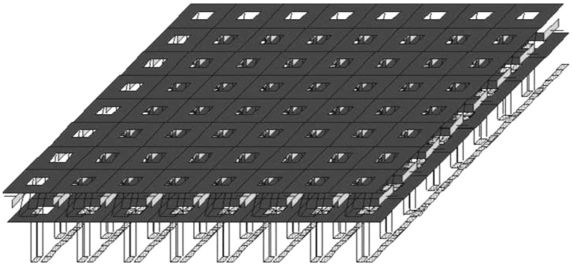
3.5 PMLP system configuration.


Electron source
Collimator lens
Aperture array
Beam blanker array
Beam deflector array
Projection lens array

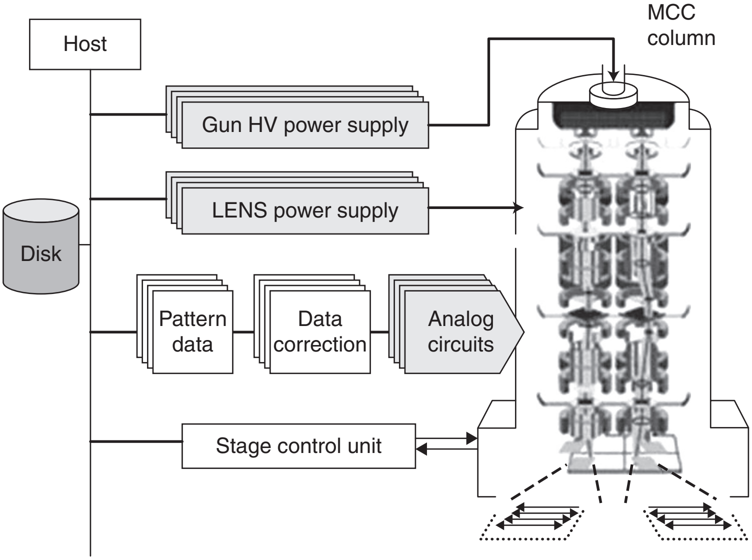
3.6 MAPPER system configuration.

and off at will. This approach is referred to as *multiple aperture pixel-by-pixel enhancement of resolution* (MAPPER).

A variant of the multiple Gaussian beam approach is to project multiple *shaped* beams. In Fig. 3.7 (Slodowski *et al.*, 2011, Doering *et al.*, 2012), an array of 64 shaped beams in a single column replaces the single shaped beam depicted in Fig. 3.2b. Introducing multiple shaped beams represents



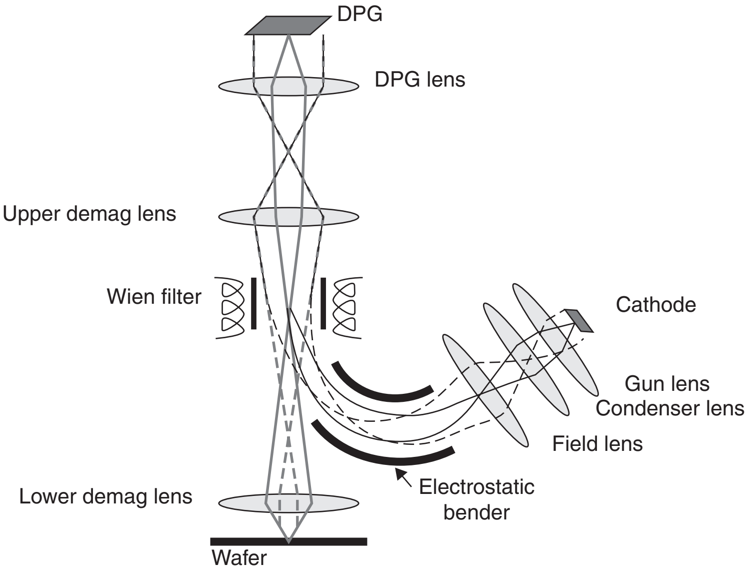
3.7 Multiple shaped beam (MSB) concept.


Host
Gun HV power supply
LENS power supply
Disk
Pattern data
Data correction
Analog circuits
Stage control unit
MCC column

3.8 Multi-column cell (MCC) system concept, schematic, with character projection (CP).

a relatively minor change to an electron column with a single shaped beam. This concept therefore builds on previously proven technology in an incremental way, thereby minimizing expense and risk.

In Fig. 3.8 (Yamada *et al.*, 2008; Yamada *et al.*, 2010; Takizawa *et al.*, 2011), multiple columns are employed, with each column having multiple cell projection capability. This concept is referred to as *multi-column cell projection* (MCC). Additional columns can be added with modest effort and expense. Both systems depicted in Figs 3.7 and 3.8 have the advantage that each beamlet has many pixels. This permits one to use relatively few beamlets to achieve high pixel parallelism. In both cases, beams can be easily added to an existing platform. This permits one to add parallelism



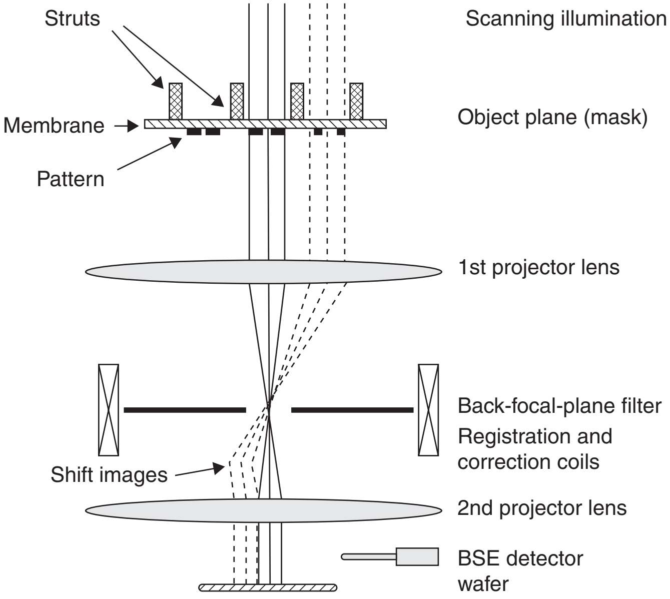
3.9 Reflective electron beam lithography (REBL) system concept.

in an incremental way, thus mitigating the risk associated with increasing system complexity.

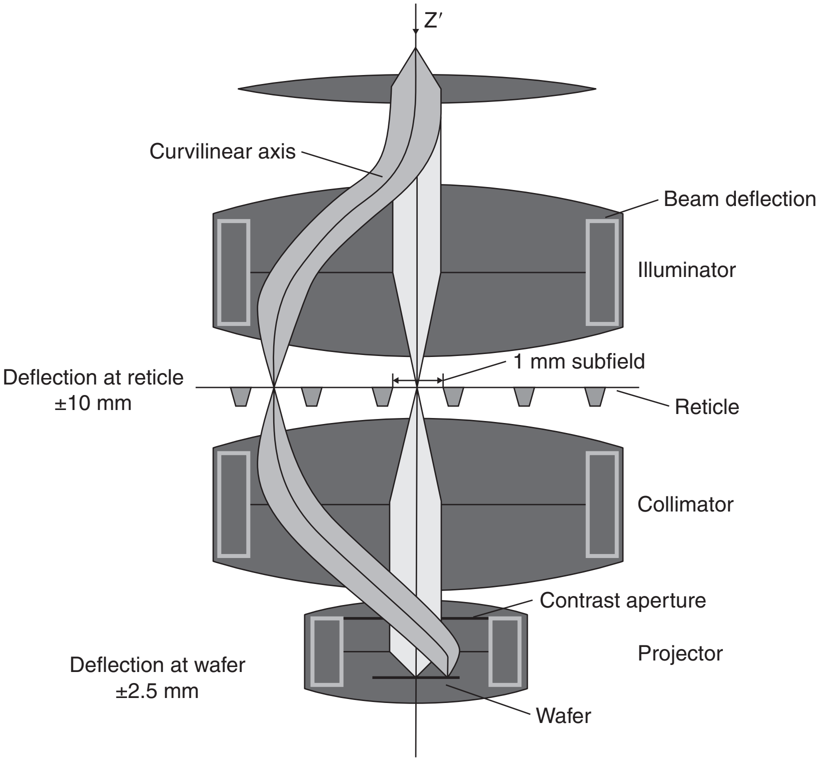
An alternative concept is shown schematically in Fig. 3.9 (McCord *et al.*, 2010). The beam from a thermionic electron source (cathode) floods a large-area *digital pattern generator* (DPG). This is an array of individual pixels, with each pixel independently addressable with a voltage. The beam is decelerated to a very low energy at the DPG, so that beam electrons are either reflected or absorbed, depending on the pixel voltage. The reflected beam containing the pattern information is then accelerated. The pixel array is demagnified onto the writing surface. The electron column is compact, thus mitigating Coulomb scattering of beam electrons.

An alternative approach is to project an electron optical image of a patterned membrane mask onto the writing surface. This approach, called *projection* e-beam lithography, forms a demagnified image of a pre-existing mask. It was first conceived by Heritage (1975) and independently by Koops and Bernhard (1975), based on the operating principle of a transmission electron microscope. Two more modern implementations are described below.

The approach in Fig. 3.10 is known as *SCattering with Angular Limitation Projection Electron Beam Lithography* (SCALPEL) (Harriott *et al.*, 1996, Stenkamp *et al.*, 2001). It was developed jointly at Bell Laboratories, Lucent Technologies, and Carl Zeiss GmbH. The approach in Fig. 3.11 is known as *Projection Reduction Exposure with Variable Axis Immersion Lenses*


Struts
Scanning illumination
Membrane
Object plane (mask)
Pattern
1st projector lens
Back-focal-plane filter
Registration and correction coils
Shift images
2nd projector lens
BSE detector
wafer

3.10 SCALPEL system concept.


Z′
Curvilinear axis
Beam deflection
Illuminator
1 mm subfield
Deflection at reticle ±10 mm
Reticle
Collimator
Contrast aperture
Projector
Deflection at wafer ±2.5 mm
Wafer

3.11 PREVAIL system concept.

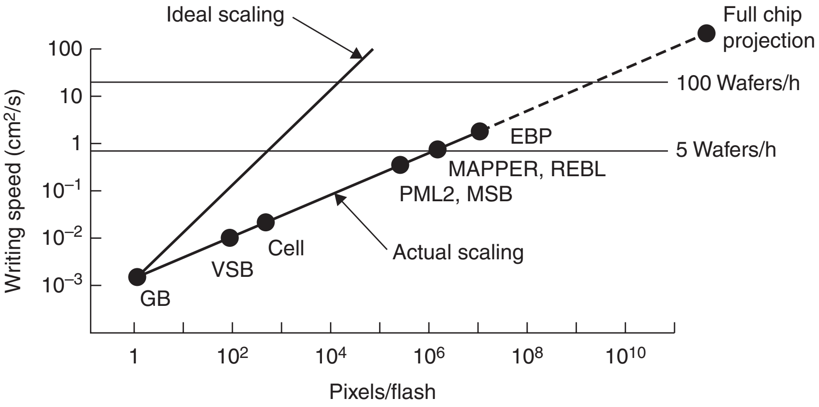
(PREVAIL) (Pfeiffer *et al.*, 1999; Golladay *et al.*, 2000; Gordon *et al.*, 2000; Pfeiffer *et al.*, 2001; Golladay *et al.*, 2001). It was developed jointly by IBM and Nikon. This latter approach features a high-emittance thermionic electron source, where emittance is defined as the product of the transverse area and the emission solid angle. Such a source has the dual advantages of uniform illumination intensity and high total current. PREVAIL also features a low-aberration deflection system, which scans the beam in a direction perpendicular to the motion of the XY stage.

Two generic approaches have been tried for the mask. In one approach, the membrane is sufficiently thin so that most of the electrons pass through unscattered. The pattern features on the mask are made of a material with high atomic number, such as tungsten or tantalum. These areas scatter the electrons out of the acceptance angle of the electron optical system, thus producing image contrast. In another approach, the mask is a stencil that transmits current only through the holes. Two complementary masks are needed in order to print all possible patterns. Pixel parallelism on the order of 10^8 pixels per flash has been demonstrated using these approaches. Indeed, electron optical imaging has been used in a variety of ways to dramatically increase pixel parallelism, thus mitigating the serial limitation.

In order for any of these methods to increase writing speed, the total writing current must increase accordingly. Ultimately, writing speed is always limited by the availability of useful writing current, regardless of the degree of pixel parallelism. The term 'useful' implies that the desired resolution is achieved. This resolution is a matter of choice (down to the ultimate limit imposed by electron scattering and secondary processes in the resist), and depends on the specific application. By degrading the resolution, the useful writing current is increased. For this reason, one tunes the resolution to the value required for the application, but no better. With the resolution chosen, one then maximizes the writing current.

An important caveat is that, in some applications, the throughput is limited by system overheads that have nothing to do with the writing current. In particular, if the pattern is relatively sparse, the system spends a small fraction of the time actually writing, and a large fraction of the time doing other tasks. These tasks include pumping the system down to adequately good vacuum, moving the mechanical stage, and waiting for the deflection fields to settle down to an appropriate level. The allocation of process time is sensitive to the particular pattern. A useful metric is the fraction of the writing area that is actually patterned. This fraction can be vanishingly small, or it can approach unity, depending on the application.

For dense patterns with a large amount of pattern data, the system spends a relatively large fraction of the time actually writing, as opposed to doing other tasks. In this case, a useful estimate of the overall throughput is the total average writing current I divided by the incident dose D in



3.12 Writing speed in units of area swept out per unit time, as a function of pixels per flash for various parallel e-beam exposure concepts.

units of charge per unit area needed for proper exposure of the resist. This ratio has units of area swept out per unit time. It represents the overall system speed at the limit where system overhead delays are insignificant. Typical e-beam writers operate in the range 0.0001–1.0 cm^2/s. By comparison, conventional optical exposure tools operate in the range 20–30 cm^2/s. Obviously, one can potentially improve e-beam writing speed by increasing the total average writing current I, decreasing the dose to the resist D (by judicious chemical engineering of the resist sensitivity), or by a combination of these.

It is interesting to plot the useful writing current as a function of pixel parallelism (measured in pixels per flash) actually achieved in practice. This is shown in Fig. 3.12 for various writing strategies described above. Ideally, one might expect the useful writing current to be proportional to the pixel parallelism, as represented by the curve labeled 'ideal scaling' in the figure. This is generally not achieved in practice, however. According to the figure, the useful writing current increases roughly as the cube root of the pixel parallelism. Again, this is due to the limitation on useful writing current imposed either by Coulomb scattering, or the ability of the electron optical system to supply the desired current. This sub-linear dependence indicates that introducing more parallelism into the exposure does not proportionally increase the writing speed.

The position of the whole curve in Fig. 3.12 along the vertical axis depends on the resolution. As one demands finer resolution, the available writing current decreases. This causes the curve to be displaced downward along the vertical axis. As one relaxes the resolution, the curve is displaced upward along the vertical axis. An estimate of the resolution is obtained from the full width at half maximum of the lateral pixel intensity distribution. In order to ensure adequate lithographic pattern fidelity, this must be a reasonably small fraction of the CD. As a rough rule of thumb, this fraction is typically chosen to be about one-fourth. A pixel resolution of 10 nm (corresponding to a CD of 40 nm) is arbitrarily assumed in the Fig. 3.12. An incident dose of 60 μC/cm^2 at 50 kV beam voltage is arbitrarily assumed as well.

The writing speed in units of area swept out per unit time is obtained by dividing the total average writing current I in the figure by the incident dose D needed to properly expose the resist. This dose is typically in the range 10–1000 μC/cm^2, depending on the choice of resist and the beam voltage. The required dose is proportional to beam energy. At higher beam voltage, on the order of 100 kV, a typical resist layer on the order of 100 nm thick is relatively transparent to the beam. In this case, a small fraction of the beam energy is deposited in the resist. Consequently, a relatively high incident dose is needed, and the writing speed is correspondingly lower. At lower beam voltage, on the order of 10 kV or less, a relatively large fraction of the beam energy is deposited in the resist. Consequently, a relatively low incident dose is needed, and the writing speed is correspondingly higher.

The useful writing current increases with increasing beam energy. This occurs because Coulomb scattering in the beam path is less pronounced at higher beam energy. Separately, the ratio of brightness to beam voltage is conserved throughout every optical system. For a given electron source, the brightness at the writing surface is therefore proportional to the beam voltage at the writing surface. These factors lead to higher writing current at higher energy, offsetting the need for higher incident dose. The writing speed (ratio I/D of current to dose) is therefore a weak function of beam energy.

As mentioned, Coulomb scattering in the electron column places a limit on useful writing current. Coulomb scattering represents a fundamental limitation, and therefore has received considerable attention (Boersch, 1954; Pfeiffer, 1971; Groves *et al.*, 1979; El-Kareh and Smither, 1979; Sasaki, 1979; Jansen, 1990; Groves, 1999; Winograd *et al.*, 1999; Han *et al.*, 2000; Kruit and Jansen, 2009) since its importance was first recognized by Boersch (1954). This scattering is more pronounced in locations within the column where the average space charge density is high. This, in turn, depends on the configuration of the electron optical column, and on operating parameters such as current, accelerating voltage, and numerical aperture. Coulomb scattering is more pronounced near crossover points in the beam, where the writing current passes through a constriction.

In the above arguments we have implicitly assumed that data rate does not limit the throughput. In other words, we have assumed that the electronic data path is able to supply pattern data quickly enough so that the beam never has to 'wait' for data. This is not always the case. We can regard the writing process as a succession of flashes, during which the beam is turned on. The frequency of flashes f is given by

$$f = \frac{I}{D \cdot A} \qquad [3.1]$$

where I is the total writing current, D is the dose required to properly expose the resist, and A is the total area exposed in one flash. The frequency f has units of s^{-1} or Hz as required. We recognize the ratio I/D as the writing speed described above, and the ratio I/A as the average current density at the writing surface. The ability of an e-beam writer to supply data at flash rate f depends on the digital data rate, the bandwidth of the analog electronics which control the beam, and the speed at which the electron optics can turn the beam on and off.

Considering the simplest case of a single Gaussian beam, we can assume some reasonable numbers. For a spot radius of 10 nm, a dose of $D = 60\ \mu C/cm^2$, and a flash rate of $f = 50$ MHz, the maximum usable writing current is about 10 nA. For writing current greater than this, the system would be limited by the data rate, rather than by usable writing current. The writing speed in this example is about $1.7 \times 10^{-4}\ cm^2/s$. At this rate it would take about 1.7 h to write an area of 1 cm^2. Several groups have reported data rates of roughly 300 MHz for Gaussian beam writers (Kelly 1981; Alles 1987; Gesley 1993). It would become difficult to achieve a data rate appreciably higher than this for an e-beam writer, owing to the necessity of synchronizing the pattern data stream with the electronics.

As a separate example, we consider a variable-shaped beam writer. Assuming a current density of 400 A/cm^2 and a dose of 20 $\mu C/cm^2$, the flash rate f must be about 20 MHz. In order to take advantage of this speed, it is necessary that other system overhead times be less than or equal to the writing time. This includes the digital-to-analog conversion (DAC), which converts the digital beam deflection coordinates to an analog signal to drive the deflector. It also includes the settling time of the analog electronics. The move and settling time for an electrostatic deflection step of a few volts is typically about 50 ns in practice. This applies to electrostatic shaping as well. This is similar to the flash time in the present example. It would become difficult to operate a single variable-shaped beam writer appreciably faster than this.

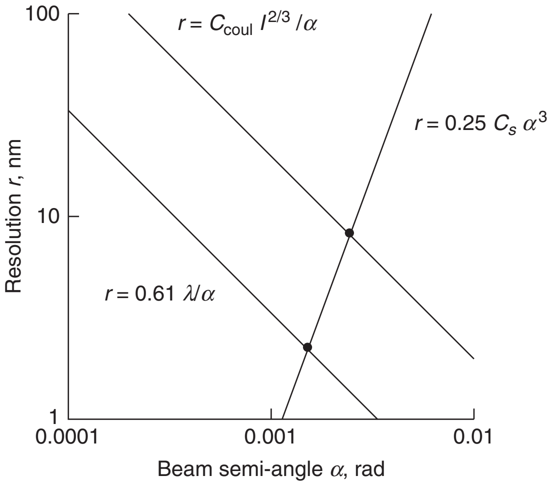
According to Equation [3.1], the flash rate f is inversely proportional to the area A exposed in a single flash. Exposing multiple pixels simultaneously, the exposure area A is the area of one pixel times the number of pixels exposed per flash. It follows that the required flash rate is inversely proportional to the pixel parallelism. The flash rate f is highest for a single Gaussian beam with one pixel per flash, and decreases for greater pixel parallelism.

3.3 The tradeoff between resolution and throughput

Given that present-day e-beam writers are slow relative to optical exposure systems, it is natural to ask whether fundamental improvement is possible. This is closely related to how the useful writing current I scales with resolution r. Electron optical resolution r depends on the choice of numerical aperture α, where α is the semi-angle of the beam cone measured in radians at the writing surface. This angle was indicated previously in Fig. 3.2. Every electron optical system (as well as every light-optical system) can be characterized by its numerical aperture. This remains true, regardless of the number of beams or the size of the projected image. The pixel resolution r is plotted as a function of beam semi-angle α in Fig. 3.13 for a typical e-beam writer. The contribution due to diffraction is given by Rayleigh's criterion, which is given for small angles α by

$$r = 0.61\frac{\lambda}{\alpha} \qquad [3.2]$$

where λ is the de Broglie electron wavelength given by



3.13 Resolution as a function of beam semi-angle. Significant contributions include Coulomb interaction and spherical aberration. Diffraction is negligible in this case.

$$\lambda = \frac{h}{p} \qquad [3.3]$$

and h is Planck's constant given by $h = 6.626 \times 10^{-34}$ Joule-s. Also p is the momentum of a beam electron.

Separately, the contribution due to spherical aberration is given by

$$r = 0.25\, C_S\, \alpha^3 \qquad [3.4]$$

where C_S is the spherical aberration coefficient of the electron optical system. These two contributions are plotted as a function of α in Fig. 3.13 for a typical e-beam writer. An optimal value of α exists where the two contributions are equal to one another. This represents the best achievable resolution. Setting the two expressions for r equal to one another and eliminating α, we obtain the optimal resolution r as follows:

$$r = 0.49\, C_S^{1/4} \lambda^{3/4} \qquad [3.5]$$

This is shown for a 50 kV beam as the lower point in the Fig. 3.13, where $\lambda = 5.36 \times 10^{-12}$ m.

The determination of resolution by diffraction and spherical aberration applies only when the writing current is sufficiently small so that Coulomb scattering is unimportant. This applies in an electron microscope, for example, where at most a single electron is present in the electron column at any given instant. As one increases the current, the Coulomb scattering dominates over diffraction as the determinant of resolution. The contribution due to Coulomb scattering can be written approximately (Jansen, 1990) by

$$r = C_{\mathrm{Coul}} \frac{I^{2/3}}{\alpha} \qquad [3.6]$$

where C_{Coul} is a proportionality constant that depends on the specific column geometry and the beam voltage, and I is the total average writing current. Applying a similar optimization argument to the above, we obtain an optimal resolution of

$$r = 0.707\, C_S^{\frac{1}{4}}\, C_{\mathrm{Coul}}^{\frac{3}{4}} I^{\frac{1}{2}} \qquad [3.7]$$

This is represented by the upper point in Fig. 3.13. To summarize, at low beam current the electron optical resolution is determined by spherical aberration and diffraction. At high beam current the electron optical resolution is determined by spherical aberration and Coulomb scattering in the beam.

On the basis of these considerations, we are now in a position to determine how the writing speed scales with resolution. First we consider an e-beam writer that is limited in useful writing current by Coulomb scattering within the beam path. From Equation [3.7], the useful writing current I can be expressed as a function of resolution r by the proportionality

$$I \propto r^2 \qquad [3.8]$$

In words, the usable current I is proportional to the square of the resolution r. Separately, the dose needed to properly expose the resist is given by

$$D \propto \frac{ne}{\pi \cdot r^2} \qquad [3.9]$$

where n is the average number of incident electrons per pixel needed to properly expose the resist, e is the electron charge, and r is the radius of the pixel. The quotient I divided by D is the area swept out per unit time by the writing beam. It represents an estimate of throughput in the limit where system overhead times are insignificant. This implies that the written pattern is relatively dense. Consequently, the system spends most of its time actually writing, as opposed to waiting. From Equations [3.8] and [3.9], this obeys the proportionality

$$\frac{I}{D} \propto r^4 \qquad [3.10]$$

This fourth-power dependence was first pointed out by Pease (2005). It is consistent with scaling trends of other forms of lithography as well (Marrian and Tennant, 2003). We deduce that the writing speed has a very strong and disadvantageous dependence on the resolution for a single-source e-beam column.

From Equation [3.10], one can improve the writing speed in principle by chemically engineering the resist to decrease the dose D needed for proper exposure. The average number of electrons needed to expose one pixel is n. The average fluctuation in this number from pixel to pixel is $\sqrt{n}$. The average fractional fluctuation ε is therefore given by $\varepsilon = \sqrt{n}/n = 1/\sqrt{n}$. This fluctuation is known as *shot noise*, and arises from the discreteness of electrons. From Equation [3.9] this places a lower limit on usable dose $D_{\min}$ given by

$$D_{\min} = \frac{e}{\pi \cdot r^2 \cdot \varepsilon^2} \qquad [3.11]$$

Assuming for example $\varepsilon = 0.02$ and $r = 20$ nm, we find in this simple model that $D_{\min} = 32\ \mu\text{C/cm}^2$. Choosing a resist with dose less than this is likely to compromise the edge smoothness of exposed features. A more detailed elaboration of this simplistic model is described by Kruit (1998), Kruit *et al.* (2004), and Kruit and Steenbrink (2005).

The useful writing current I depends on the beam voltage, column length, and the numerical aperture α of the electron optical column, and is therefore very sensitive to the specific column configuration. The various configurations described above have very different values of the maximum usable current. Unfortunately, no general theory exists that allows prediction of the maximum current for all systems. In practice, one must rely on simulation and measurement to determine this. Monte Carlo simulation gives an accurate prediction of resolution for a given system configuration and writing current. It is also computationally intensive and time-consuming. Typically, one uses Monte Carlo simulation to analyze a single operating point. This consists of finding the resolution for a specific combination of electron column configuration, beam voltage, beam writing current, and beam semi-angle α. Given this computed resolution, one can easily explore perturbations about this base operating point using known trend formulas (Jansen, 1990; Kruit and Jansen, 2009). For a column designer this predictive exercise is essential (Orloff, 2009; Munro, 2012), since the actual fabrication and testing of the column is expensive and time-consuming. One would like to try as hard as possible to get it right the first time. Despite one's best efforts, this is very difficult to achieve in practice.

The above analysis rests on the assumption that the maximum usable writing current is limited by Coulomb scattering in the beam. This can only occur only when the electron source is able to supply sufficient current. This does not always occur in practice. There are two generic reasons for this. First, it is virtually impossible to use all of the source current to form the writing beam; i.e., one inevitably throws away a significant fraction of the emission using the various apertures needed to achieve proper illumination and resolution. Second, the specific brightness obtainable at the writing surface cannot exceed that of the source. Brightness B is defined at the writing surface for small angles α as

$$B = \frac{I}{(\pi \cdot r^2)(\pi \cdot \alpha^2)} \qquad [3.12]$$

This is the current per unit transverse area per unit solid angle. The specific brightness is defined here as the ratio of the brightness B divided by the beam voltage V. This ratio is conserved in every charged particle optics system. Mathematically,

$$\frac{B}{V} = const \qquad [3.13]$$

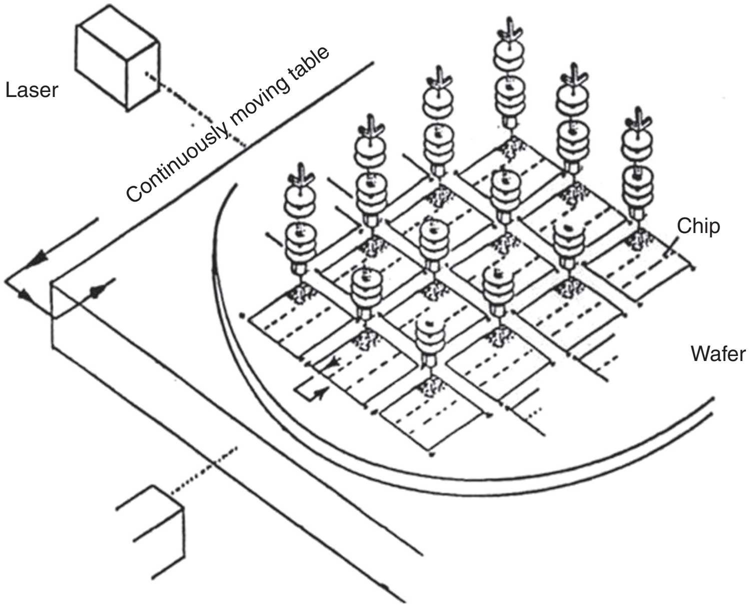
It follows that the specific brightness at the writing surface can never exceed that of the source. This imposes a fundamental limit on available current. In the event that the source cannot supply enough current for the Coulomb interaction to be a limit, the writing speed can only be lower than the Coulomb-limited case.

3.4 Distributed systems

These arguments prompt us to ask whether there is any possible way around the adverse scaling of writing speed with resolution. A possible answer lies in the fact that, for all of the systems described above, we have made the implicit assumption that the e-beam writer has a single electron source, a single electron column, or both. This has been true to this point regardless of the degree of pixel parallelism. This general approach has the multiple advantages of simplicity, reliability, and low cost, but the disadvantage of low throughput.

It is possible in principle to replicate any of the systems previously described. This would of course require multiple electron sources and multiple electron columns. One would only need to scale the number of systems with the fourth power of the resolution. This does not represent a fundamental limitation, but only one of practicality and expense. Following this reasoning, one could imagine a system in which the beam current is spread over a large volume, so that Coulomb scattering is unimportant. At the same time, one can envision multiple electron sources, so that the ability of the electron optical system to supply current is increased accordingly. We call such a system a *distributed* system.

One such proposal uses many miniature electron columns, each with its own source. This concept is shown schematically in Fig. 3.14 (Chang *et al.*, 1996, 2001). Each column is potentially inexpensive, owing to its small size and relative simplicity. In principle, one could pack a great many columns

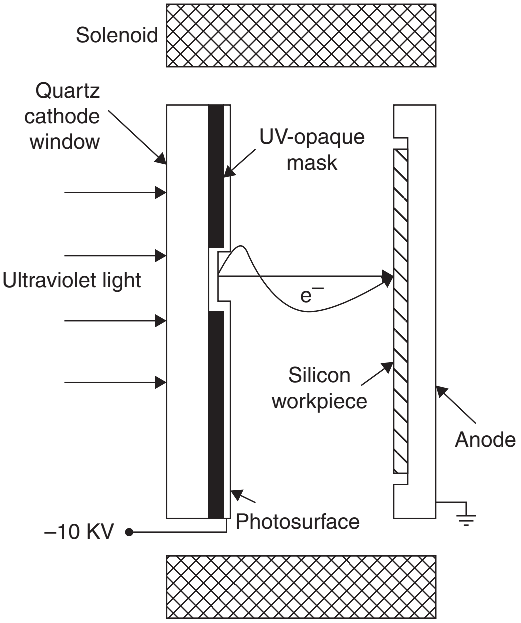


3.14 Distributed system with multiple, arrayed electron columns.

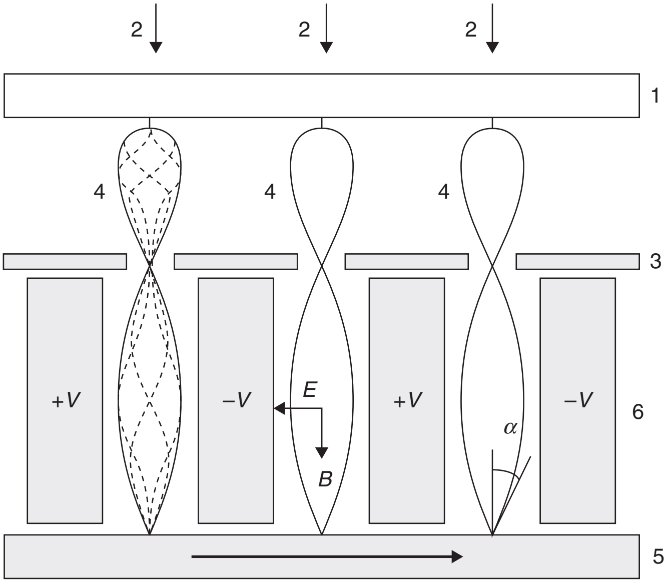
into a modest area, thereby enabling high throughput. This column concept has already been proven to provide high lithographic resolution.

Another distributed system was actually proposed in the late 1960s, and is shown schematically in Fig. 3.15. It is called *electron lithography image projection system* or ELIPS (O'Keefe and Handy, 1969; Malmberg *et al.*, 1973). A large-area, flat photocathode is patterned with the desired pattern to be printed. This cathode is placed opposite the writing surface, with the cathode at a high negative potential relative to the writing surface. The entire apparatus is immersed in the uniform magnetic field of a solenoid. The magnetic field is oriented perpendicular to the photocathode and writing surface. Illuminating the cathode from the backside with ultraviolet light causes photoelectron emission from those pattern areas that are not blocked by an opaque material. A one-to-one image of the photocathode is transferred to the writing surface. The writing current is distributed over a large volume with no intermediate crossover. Coulomb scattering is unimportant in this system. Because of the large area, a large writing current can be obtained, even for low emission current density. This system has enormous pixel parallelism, but does not suffer from the writing speed limitations of the other systems described previously.

A new patterned photocathode is needed for every distinct pattern. This practical limitation can be overcome in principle by considering a variant of the ELIPS concept shown schematically in Fig. 3.16 (Groves and


Solenoid
Quartz cathode window
UV-opaque mask
Ultraviolet light
e⁻
Silicon workpiece
Anode
Photosurface
−10 KV

3.15 ELIPS system concept.


2
2
2
1
4
4
4
3
+V
−V
+V
−V
E
B
α
6
5

3.16 Distributed fixed-axis (DIFA) system configuration.

Kendall, 1998, Pickard *et al.,* 2003). A transparent plate (1) is patterned with an array of point-like photo-emitters. The emitters are illuminated from the top of the figure by an array of focused light beams (2). The bottom surface of the plate (1) is covered with a non-emissive metal film, and is elevated to a high negative potential relative to the rest of the system, which is nominally at ground potential. The photo-emitted electrons are accelerated to an anode (3), which is at ground potential, forming an array of beams (4). The entire apparatus is immersed in the uniform field of a solenoid (not shown in Fig. 3.16) with the magnetic field B oriented parallel to the beam axes. Each beam forms a 1:1 image of the source, and is allowed to pass through a small aperture in the anode (3). Each beam drifts to the writing surface (5), where a second 1:1 image of the source is formed. The leftmost beam in the figure shows the individual electron trajectories spiraling in the magnetic field. The beam semi-angle α is given by the ratio of transverse to axial velocity components. Assuming the transverse velocity is thermal, this semi-angle is naturally a few milliradians, with no physical aperture needed. This permits resolution on the order of a few nanometers. A set of parallel plates (6) extend a long distance into the plane of the figure. The plates are excited to a potential $+V$ or $-V$, with neighboring plates having opposite sign of the potential. This creates a uniform electric field E in the transverse direction. Together with the uniform axial magnetic field B, the beams are deflected. This deflector has excellent electron optical properties (Kuo and Groves, 1983). Owing to the magnetic field, the net deflection in the writing plane is into or out of the plane of the figure, depending on the direction of the electric field vector. The beams intersect the writing surface at normal incidence. The deflection aberration is exceedingly low. By ramping the voltage V, the beams scan in the direction into or out of the plane of the figure. Many beams can be added along this axis, creating a two-dimensional array. By mechanically scanning the writing surface (5) in the direction of the arrow, and blanking each beam individually, an arbitrary pattern can be created on the writing surface. The beams can be individually blanked and unblanked by turning the light source off and on. Fine-line lithography has been demonstrated in the laboratory at low throughput with this concept. Successful large-scale implementation will rely on a nanoscale photoelectron source being identified and demonstrated.

In summary, distributed system concepts in principle circumvent the limitation on useful writing current due to Coulomb scattering and source current limitation. This comes with the additional complexity of multiple sources. In principle, all of the single-source options described above can be replicated to form distributed systems. The use of e-beam direct write lithography in high-volume semiconductor chip manufacturing is of renewed interest (Lin, 2012), as the cost of conventional optical lithography escalates.

3.5 Ultimate lithographic resolution

Ultimate lithographic resolution is limited by interaction of the e-beam with the recording medium. Every incident beam electron undergoes scattering in the resist layer. There are two generic scattering processes: *inelastic* and *elastic* scattering. In inelastic scattering, the incident electron transfers some of its energy to the resist. This is responsible for proper exposure of the resist. Inelastic scattering is characterized by negligibly small scattering angle. In elastic scattering the incident electron is scattered by a target atom nucleus, with negligible transfer of energy. Elastic scattering is characterized by relatively large scattering angle. This gives rise to spreading of the beam in the target material.

Scattering in the recording medium and substrate is characterized by a *mean free path*. This is the average distance the electron travels in the material between scattering events. The mean free path μ_e for elastic scattering is given by

$$\mu_e = \frac{1}{N\,\sigma_e} \qquad [3.14]$$

where N is the number of target atoms per unit volume, and σ_e is the *total cross-section* for elastic scattering. The number N is given by

$$N = \frac{N_0 \rho}{A} \qquad [3.15]$$

where N_0 is Avogadro's number given by $N_0 = 6.0221 \times 10^{23}$, ρ is the mass density, and A is the atomic number. The total cross-section σ_e is the probability per unit incident intensity that a scattering event will occur from a single target atom. It has units of area, and is given in SI units approximately by

$$\sigma_e = \frac{Z^{4/3} h^2}{2\pi \cdot m \cdot E} \qquad [3.16]$$

where Z is the atomic number of the target material, h is Planck's constant given by $h = 6.6261 \times 10^{-34}$ J.s, m is the mass of the electron given by $m = 9.1094 \times 10^{-31}$ kg, and E is the kinetic energy of the incident electron in Joules. Inserting the values of the constants, this leads to the practical expression

$$\sigma_e = (4.788 \times 10^{-19}) \frac{Z^{4/3}}{E_{\mathrm{eV}}} \qquad [3.17]$$

in units of m^2, where E_{eV} is the kinetic energy of the incident electron in electron volts. As an example, inserting the constants for silicon, $\rho = 2.34$ g/cm^2, $A = 28$, $Z = 14$, we find that the mean free path of a 50 keV electron is 62 nm in silicon. Finally, the total cross-section for inelastic scattering is related to elastic scattering approximately by

$$\frac{\sigma_i}{\sigma_e} \approx \frac{10}{Z} \qquad [3.18]$$

The ratio of the mean free path for μ_{i} for inelastic scattering divided by the mean free path μ_{e} for elastic scattering is the reciprocal of this. For a 50 kV electron in silicon, the inelastic mean free path is about 87 nm. As the beam penetrates the substrate material, it continually loses energy. Consequently, the mean free path constantly decreases. In summary, the higher the instantaneous energy, the more transparent the material becomes.

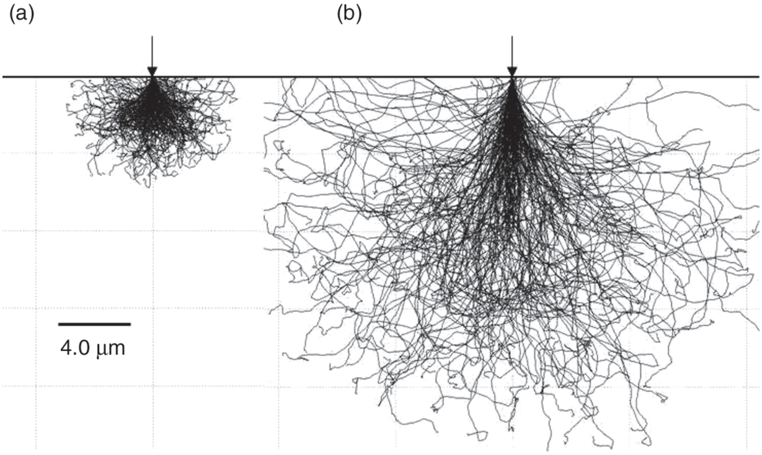
Most of the beam electrons pass through the resist layer and into the substrate material. There the electrons are multiply scattered, both elastically and inelastically. This is shown in Fig. 3.17 for bulk silicon using the Casino Monte Carlo simulation program (Drouin *et al.*, 2007). Elastic scattering causes the beam to broaden laterally. Some fraction of the beam is backscattered, and exits through the resist layer, causing unwanted exposure. Inelastic scattering causes the beam to deposit energy in the substrate as heat (Groves, 1996). This beam-induced heating can alter the resist sensitivity.

On average the beam electrons travel to a depth R into the substrate, where the range β is given in micrometers approximately by (Gruen, 1957):

$$\beta_{\mu\mathrm{m}} = 0.046 \frac{(V_{\mathrm{kV}})^{1.75}}{\rho_{\mathrm{g/cm^3}}} \qquad [3.19]$$

where V_{kV} is the beam voltage in kilovolts and $\rho_{\mathrm{g/cm3}}$ is the density of the substrate material in g/cm^3. The range β represents an estimate of the lateral distance in micrometers over which the backscattered current causes unwanted exposure of the resist. The strong dependence on incident beam energy is clearly seen in Fig. 3.17.

This stray exposure is small for isolated pattern features, and relatively large for dense pattern features. This causes the amount of stray exposure to depend on the specific pattern being written. This is known as *proximity effect*. It is possible to correct for proximity effect by applying greater incident dose to isolated features and less incident dose to dense features (Parikh and Kyser, 1979; Rishton and Kern, 1987).



3.17 Monte Carlo simulations of individual electron trajectories in bulk silicon, (a) 25 kV incident beam voltage, (b) 50 kV incident beam voltage.

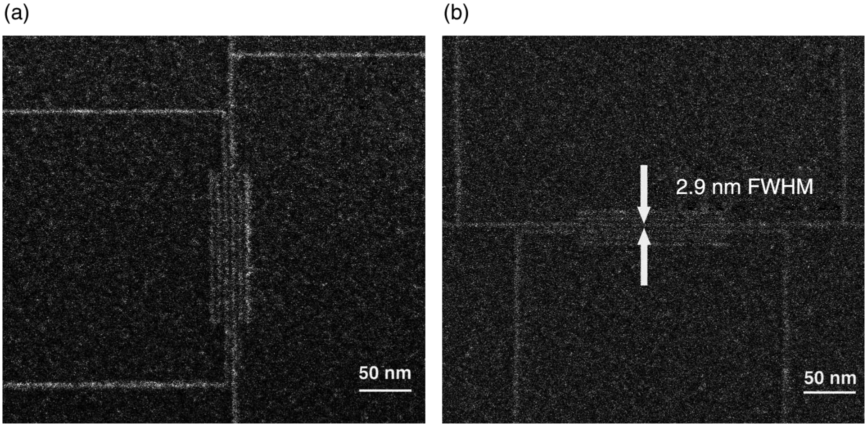
The exposure $E(r)$ as a function of two-dimensional position r in the writing plane can be expressed in terms of the incident dose $D(r)$ as a convolution as follows (Pavkovich, 1986):

$$E(r) = \int D(r')\,\rho(r - r')\,\mathrm{d}^2 r' \qquad [3.20]$$

where $\rho\ (r)$ is known as the *point spread function*. Physically, this function represents the blurring caused by electron scattering at any given point in the pattern. The function ρ (r) in turn can be approximated by

$$\rho(r) = \frac{1}{\pi(1+\eta)}\left[\frac{1}{\alpha^2}\exp\left(\frac{-r^2}{\alpha^2}\right) + \frac{\eta}{\beta^2}\exp\left(\frac{-r^2}{\beta^2}\right)\right] \qquad [3.21]$$

where α is the range of the forward scattering in the resist layer, β is the range of the backscattering from the substrate given above, and η is the fraction of the total incident current which is backscattered. Typically the forward scattering range α is of the order of a few tens of nanometers, while the backscatter range β is of the order of a few micrometers to a few tens of micrometers, depending on the incident beam voltage. The fraction η is typically between 0.2 and 0.6, depending on the beam voltage and substrate material. The reader can verify that the integral of ρ (r) over all transverse coordinates is unity as required. In order to correct for proximity effect, we must calculate the incident dose D(r) required to produce an ideal desired



3.18 Ultimate lithographic resolution using EBID technique.

exposure E(r). Mathematically, this amounts to solving the above integral equation for D(r), assuming the other quantities are known. Software is commercially available to accomplish this.

Ultimate lithographic resolution is limited by scattering and secondary processes within the resist layer. The resist layer thickness is typically of the order of 100 nm or less. Elastic scattering causes the beam to broaden laterally in the resist layer. Generally, this broadening is inversely related to the incident beam energy. Secondary processes that give rise to exposure of the resist generally have a finite range, also contributing to line broadening. Ultimate lithographic resolution therefore depends on the process, rather than on the electron optical resolution of the beam writer.

An intriguing way to extend lithographic resolution beyond what is possible using conventional patterning in resist is *E-Beam-Induced Deposition* (EBID). In this process, a low partial pressure of a precursor gas is intentionally introduced into the chamber of a conventional scanning electron microscope (SEM). The focused e-beam induces the material from the gas to be condensed onto the writing substrate in solid form. The resulting lithographic resolution is very high. An example is shown in Fig. 3.18, which shows lithographic details approximately three nanometers in width (van Oven, *et al.,* 2011). A platinum-bearing precursor gas used was used. Ultimate lithographic resolution depends in general on the recording medium. Ongoing research leads to continuing improvement.

3.6 Electron-beam patterning of photomasks for optical lithography

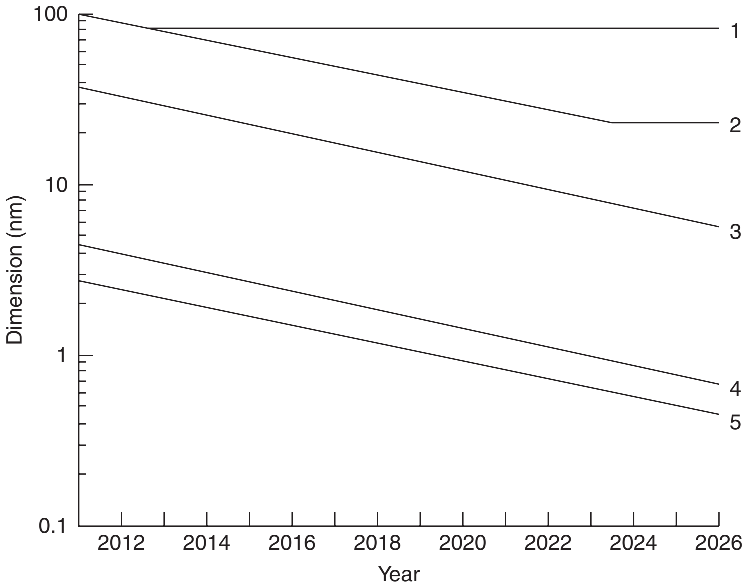
Mainstream high-volume manufacturing of integrated circuit chips relies on light-optical lithography. This in turn relies on a pre-existing mask that contains the complete pattern for a single exposure level, also known as a *mask level*. The pattern on the mask is optically reduced by 4× onto the resist-coated wafer, with the exposure repeated once for each chip on the wafer. The standard mask size is 150 mm × 150 mm in size, and the mask is typically made of quartz. The exposure light source is an ArF excimer laser operating at 193 nm wavelength. Integrated circuit chips typically have between 30 and 45 mask levels, each requiring a different patterned mask.

E-beam lithography is the method of choice for patterning photomasks in high-volume manufacturing of VLSI chips. This takes advantage of the inherent high resolution and pattern-generating capability. Historically, the limited throughput of e-beam lithography has not been a serious limitation, because the writing time is amortized over all of the wafers printed with that mask. Present-day mask writing time can be as long as a few tens of hours for the most complex mask levels, and under one hour for the simplest mask levels.

The fidelity with which the printed pattern on the wafer approximates the original design is crucially important. Broadly, pattern fidelity consists of accurate placement of pattern features over the lithographic area, and precise control of printed feature size and quality.

The mask levels must overlay relative to one another to a precision that is a reasonably small fraction of the minimum feature size or Critical Dimension (CD). The CD is nominally the half-pitch of dense pattern features on the wafer. For 4× reduction, the CD on the mask is obviously four times larger than on the wafer. The mask contains an enormous number of pattern features. As an example, a CD of 22 nm on the wafer requires a CD of 88 nm on the mask. A chip size of 26 mm × 33 mm on the wafer requires a patterned area of 104 mm × 132 mm on the mask. Dividing the chip area by the square of the CD yields an upper limit of 1.8×10^{12} pattern features in either case.

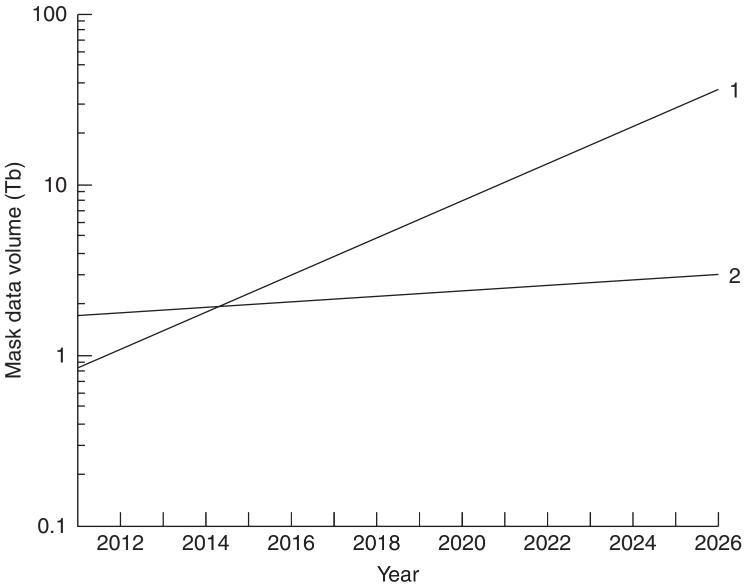
The specifications governing pattern fidelity are detailed in the International Technology Roadmap for Semiconductors (ITRS). This roadmap is available at ITRS (2011) in the references. Substantial background information about manufacturing of state-of-the-art electronic chips is given. The continuing economic success of this industry relies on a business model of continuous performance improvement, while remaining affordable to the consumer. Consequently, critical specifications must be constantly tightened over time. Five critical lithographic specifications are plotted in Fig. 3.19 as a function of year, based on estimates taken directly from the ITRS Roadmap 2011 version (ITRS Roadmap (2011), http://www.itrs.net/.). Innovation must continue for the e-beam mask writer, the associated process, and the metrology in order to achieve these specifications. In many cases, no known solution exists for future years, but solutions must be found, in order for the trend to continue as it has in the past.



3.19 Critical dimensional specifications for photomasks as a function of year of introduction in high-volume manufacturing, according to the ITRS.

With each chip generation, more pattern features must be printed on the mask. This in turn requires higher pattern data volume. The uncompressed pattern data volume in Tb is shown as a function of year in Fig. 3.20 for EUV masks and optical masks. This is based on estimates taken directly from the ITRS roadmap 2011 version. The electronic data path must be improved with each generation, in order to process the increasing pattern data volume at a sufficiently high rate. This includes translating the design data into a form which is usable by the e-beam writer, while adding various corrections for inherent errors in pattern fidelity. It also includes the step in which the translated pattern data are fed to the e-beam during writing.

The increasing number of pattern features, together with the tightening lithographic performance specifications, cause the mask writing time to increase. Present-day masks are written using variable-shaped beam writers with a spot current density of 400 A/cm^2. This tests the limit of useful writing current, as well as the achievable data rate for a single beam system. Innovation is required for the e-beam mask writer in order to achieve a reasonable writing time on the order of 10 h per mask in the future. Multiple beam systems show promise, but require significant financial investment in order to demonstrate all performance specifications by the time they are needed in manufacturing.



3.20 Mask data volume as a function of year of introduction in high-volume manufacturing, according to the ITRS.

A potentially significant application of e-beam written masks is nanoimprint lithography (Kumar and Whitesides, 1993; Haisma *et al.,* 1996; Chou *et al.,* 1996; Chou *et al.,* 2002). This subject is reviewed in detail in the Chapter 9 by D. Resnick.

3.7 Conclusion

A focused e-beam represents the smallest, finest practical writing pencil known, with the capability of producing pattern features down to a few nanometers in size. This is roughly an order of magnitude smaller than mainstream optical lithography. E-beam lithography does not rely on a pre-existing, patterned mask, but can write the pattern directly from data stored on a computer. For these reasons, e-beam lithography remains the method of choice for fabricating nanometer-scale structures in low volume.

The historical Achilles heel of e-beam lithography has been its low throughput. This results from the serial nature of the exposure. This is in contrast to optical lithography, in which the entire pattern is exposed in a single sweep. The low throughput also results from the limitation on useful writing current imposed by the electron source, or alternatively by Coulomb scattering of electrons in the beam path.

Low throughput can be mitigated by exposing many pixels in parallel. For a single electron source, the area swept out per unit time is proportional to the fourth power of the resolution. This is true regardless of the pixel parallelism, as long as a single electron source is used. Consequently, the only hope for fundamentally improving throughput is to use a distributed system with multiple electron sources. Such system concepts exist, and are presently the subject of intensive research and development.

3.8 Acknowledgements

The author gratefully acknowledges the College of Nanoscale Science and Engineering, University at Albany, State University of New York, and Agreement Number W15QKN-07–9-0001 between the ARDEC Benet Laboratories and the Arsenal Business and Technology Partnership for funding to support the writing of this book chapter.

3.9 References

Abboud F, Dean R, Doering J, Eckes W, Gesley M, Hofmann U, Mulera T, Naber R, Pastor M, Phillips W, Raphael J, Raymond F and Sauer C (1997), Multipass gray printing for new MEBES 4500S mask lithography system, Proc. *SPIE*, **3096**, 116–119.

Alles D, Biddick C, Bruning J, Clemens J, Collier R, Gere E, Harriott L, Leone F, Liu R, Mulrooney T, Nielsen R, Paras N, Richman R, Rose C, Rosenfeld D, Smith D and Thomson M (1987), EBES4: A new electron-beam exposure system, *J. Vac. Sci. Technol. B*, **5**(1), 47–52.

Boersch H, *Zeitschrift f. Physik*, **139**, 115–125.

Buck D and Shoulders K (1958), An approach to microminiature printed systems, *Proc. Eastern Joint Computer Conf., Philadelphia, PA, 3–5 December*, American Institute of Electrical Engineers, New York, 55–59.

Chang T, Wilson A, Speth A and Ting C (1976), *Proc. Seventh Int. Conf. on Electron and Ion Beam Sci. Tech.*, R. Bakish (ed.), Electrochemical Soc., Princeton, NJ, 392.

Chang T (1975), *J. Vac. Sci. Technol.*, **12**, 1271.

Chang T, Thomson M, Kratschmer E, Kim H, Yu M, Lee K, Rishton S and Hussey B (1996), Electron beam microcolumns for lithography and related applications, *J. Vac. Sci. Technol B.*, **14**(6), 3774–3781.

Chang T, Mankos M, Lee K and Muray L (2001), Multiple electron beam lithography, *Microelectron. Eng.*, **57–58**, 117–135.

Chou S, Krauss P and Renstrom P (1996), Nanoimprint lithography, *J. Vac. Sci. Technol. B*, **14**(6), 4129–4133.

Chou S, Keimel C and Gu J (2002), Ultra fast and direct imprint of nanostructures in silicon, *Nature*, **417**, 835–837.

Doering H-J, Elster T, Klein M, Heinitz J, Schneider M, U. Weidenmueller U, M. Slodowski M, I.A. Stolberg I and Dorl W (2012), Optimization of MSB for future technology nodes, *Proceedings of SPIE*, **8323**, 83232D.

Drouin D, Couture A, Joly D, Tastet X, Aimez V and Gauvin R (2007), CASINO V2.42—a fast and easy-to-use modeling tool for scanning electron microscopy and microanalysis users, *Scanning*, **29**, 92–101.

El-Kareh A and Smither M (1979), *J. Appl. Phys*., **50**, 5596–5599.

Freed R, Gubiotti T, Sun J, Kidwingira F, Yang J, Ummethala U, Hale L, Hench J, Kojima S, Mieher W, Bevis C, Lin S-J and Wang W-C (2012), Reflective electron-beam lithography: progress toward high-throughput production capability, *Proc. SPIE*, **8323**, 83230H.

Gesley M, Colby D, Raymond F, McClure D and Abboud F (1993), Electrodynamics of fast beam blankers, *J. Vac. Sci. Technol. B*, **11**(6), 2378–2385.

Golladay S, Pfeiffer H, Rockrohr J and Stickel W (2000), PREVAIL alpha system: status and design considerations, *J. Vac. Sci. Technol. B*, **18**(6), 3072–3078.

Golladay S, *et al*. (2001), PREVAIL-EPL alpha tool: early results, *J. Vac. Sci. Technol. B*, **19**(6), 2459–2467.

Gordon M, Enichen W, Golladay S, Pfeiffer H, Robinson C and Stickel W (2000), PREVAIL: dynamic correction of aberrations, *J. Vac. Sci. Technol. B*, **18**(6), 3079–3083.

Goto E, Soma T and Idesawa M (1978), Design of a variable-aperture projection and scanning system for electron beam, *J. Vac. Sci. Technol*., **15**, 883–886.

Groves T, Hammond D and Kuo H-P (1979), Electron beam broadening effects caused by discreteness of space charge, *J. Vac. Sci. Technol.*, **16**(6), 1680–1684.

Groves T (1996), Theory of beam-induced substrate heating, *J. Vac. Sci. Technol. B*, **14**(6), 3839–3844.

Groves T and Kendall R (1998), Distributed multiple variable shaped electron beam column for high throughput maskless lithography, *J. Vac. Sci. Technol. B*, **16**(6), 3168–3173.

Groves T (1999), Theory of Coulomb scattering in particle beams using Markov's method, *J. Vac. Sci. Technol. B*, **17**(6), 2808–2813.

Gruen A (1957), Lumineszenz-photometrische Messungen der Energieabsorption im Strahlungsfeld von Elektronenquellen Eindimensioner Fall in Luft, *Zeitschrift f. Naturforschung*, Band 12a, Heft 2, 89–95.

Hagen C, Grigorescu A, van Dorp W and Kruit P (2010), *Micro- and Nano-Engineering*, Elsevier.

Haisma J, *et al.* (1996), Mold-assisted nanolithography: a process for reliable pattern replication, *J. Vac. Sci. Technol*. B, **14**(6), 4124–4128.

Harriott L, Berger S, Biddick C, Blakey M, Bowler S, Brady K, Camarda R, Connelly W, Crorken A, Custy J, Dimarco R, Farrow R, Felker J, Fetter L, Freeman R, Hopkins L, Huggins H, Knurek C, Kraus J, Liddle J, Mkrtychan M, Novembre A, Peabody M, Tarascon R, Wade H, Waskiewicz W, Watson G, Werder K and Windt D (1996), Preliminary results from a prototype projection electron-beam stepper: scattering with angular limitation projecton electron beam lithography proof-of-concept system, *J. Vac. Sci. Technol. B*, **14**(6), 3825–3833.

Han L, Pease R, Meisburger W, Winograd G and Takahashi K (2000), Scaled measurements of global space charge induced image blur in electron beam projection system, *J. Vac. Sci. Technol*. B, **18**(6), 2999–3003.

Heritage M (1975), Electron-projection microfabrication system, *J. Vac. Sci. Technol*., **12**(6), 1135–1140.

Herriott D, *et al.* (1975), EBES, a practical electron lithographic system, *IEEE Trans. Electron Devices*, **21**, 385–392.

Jansen G (1990), Coulomb interactions in particle beams, *Adv. Electronics Electron Phys*., Suppl. **21**, Academic Press, New York.

Kelly J, Groves T and Kuo H-P (1981), A high-current, high-speed electron beam lithography column, *J. Vac. Sci. Technol*., **19**(4), 936–940.

Klein C, Klikovits J, Loeschner H and Platzgummer E (2011), eMET: 50 keV electron multibeam mask exposure tool, *Proc. SPIE*, **7970**, 797011.

Koops H and Bernhard W (1975), Electron-beam projection systems with compensated chromatic field aberrations, *J. Vac. Sci. Technol*., **12**(6), 1141–1145.

Kruit P (1998), High throughput electron lithography with the multiple aperture pixel-by-pixel enhancement of resolution concept, *J. Vac. Sci. Technol. B*, **16**(6), 3177–3180.

Kruit P, Steenbrink S, Jager R and Wieland M (2004), Optimum dose for shot noise limited CD uniformity, *J. Vac. Sci. Technol. B*, **22**(6), 2948–2955.

Kruit P and Steenbrink S (2005), Local critical dimension variation from shot-noise related line-edge roughness, *J. Vac. Sci. Technol. B*, **23**(6), 3033–3036.

Kruit P, Steenbrink S and Wieland M (2006), Predicted effect of shot noise on contact hole dimension in e-beam lithography, *J. Vac. Sci. Technol. B*, **24**(6), 2931–2935.

Kruit P and Jansen G (2009), Space charge and statistical Coulomb effects, In: J. Orloff (ed.), *Handbook of Charged Particle Optics*, Second Edition, CRC Press, 341–389.

Kumar A and Whitesides G (1993), Features of gold having micrometer to centimeter dimensions can be formed through a combination of stamping with an elastomeric stamp and an alkanathiol ink followed by chemical etching, *Appl. Phys. Letters*, **63**(14), 2002–2004.

Kuo H-P and Groves T (1983), A large-area deflection system with very low aberration, *J. Vac. Sci. Technol. B*, **1**(4), 1316–1321.

Lin B (2012), Future of multiple e-beam direct write systems, *Proc. SPIE*, **8323**, 832302.

Malmberg P, O'Keefe T and Sopira M (1973), *J. Vac. Sci. Technol*., **10**(6).

Marrian C and Tennant D (2003), Nanofabrication, *J. Vac. Sci. Technol. A*, **21**(5), S207–S215.

Moellenstedt G and Speidel R (1960), *Physikalsche Blaetter,* **16**, 192–198.

Munro's Electron Beam Software, Ltd. (2012), http://www.mebs.co.uk/

O'Keefe T and Handy R (1969), An electron imaging system for the fabrication of integrated circuits, *Solid State Electronics*, **12**, 841–855.

Orloff J (ed.) (2009), *Handbook of Charged Particle Optics*, Second edition, CRC Press.

Parikh M and Kyser D (1979), *J. Appl. Phys*., **50**, 1104.

Pavkovich J (1986), Proximity effect correction calculations by the integral equation approximate solution method, *J. Vac. Sci. Technol. B*, **4**(1), 159–163.

Pease R (2005), Maskless Lithography, *Microelectronic Engineering*, **78–79**, 381–392.

Pease R and Chou S (2008), Lithography and other patterning techniques for future electronics, *Proc. IEEE*, **96**(2), 262–270.

Pease R (2010), To charge or not to charge: 50 years of lithographic choices, *J. Vac. Sci. Technol*., B **28**(6), C6A1–C6A6.

Petric P, Bevis C, McCord M, Carroll A, Brodie A, Ummethala U, Grella L, Cheung A and Freed R (2010), Reflective electron beam lithography: a maskless ebeam direct write lithography approach using the reflective electron beam lithography concept, *J. Vac. Sci. Technol. B*, **28**(6), C6C6–C6C13.

Pfeiffer H and Loeffler K (1970), A high-current square-spot probe for micro-pattern generation, *Proc. Soc. De Francaise, Microscopie Electronique*.

Pfeiffer H (1971), *Eleventh Symposium on Electron, Ion, and Laser Beam Technology*, 239–242.

Pfeiffer H (1978), Variable spot shaping for electron-beam lithography, *J. Vac. Sci. Technol.*, **15**(3), 887–890.

Pfeiffer H (1979), Character projection in high-throughput e-beam lithography, *IEEE Trans. Electron Dev.*, **ED-26**, 663–667.

Pfeiffer H, Groves T and Newman T (1988), High Throughput, High Resolution Electron Beam Lithography, *IBM J. Research*, **32**(4), 494–501.

Pfeiffer H, Dhaliwal R, Golladay S, Doran S, Gordon M, Groves T, Kendall R, Lieberman J, Petric P, Pinckney D, Quickle R, Robinson C, Rockrohr J, Senesi J, Stickle W, Tressler E, Tanimoto A, Yamaguchi T, Okamoto K, Suzuki K, Okino T, Kawata S, Morita K, Suziki S, Shimizu H, Kojima S, Varnell G, Novak W, Stumbo D and Sogard M (1999), Projection reduction exposure with variable axis immersion lenses: next generation lithography, *J. Vac. Sci. Technol. B*, **17**(6), 2840–2846.

Pfeiffer H, Dhaliwal R, Golladay S, Doran S, Gordon M, Kendall R, Lieberman J, Pinckney D, Quickle R, Robinson C, Rockrohr J, Stickel W, Tressler E, Tanimoto A, Yamaguchi T, Okamoto K, Suzuki K, Miura T, Okino T, Kawata S, Morita K, Suziki S, Shimizu H, Kojima S, Varnell G, Novak W and Sogard M (2001), PREVAIL e-beam stepper tool, *Microelectron. Eng.*, **57–58**, 163–172.

Pickard D, Groves T, Meisburger W, Crane T and Pease R (2003), Distributed axis electron beam technology for maskless lithography and defect inspection, *J. Vac. Sci. Technol. B*, **21**(6), 2834–2838.

Platzgummer E, Loeschner H and Gross G (2008), Projection maskless patterning for nanotechnology applications, *J. Vac. Sci. Technol. B*, **26**(6), 2059–2063.

Rishton S and Kern D (1987), Point exposure distribution measurements for proximity correction in electron beam lithography on a sub-100 nm scale, *J. Vac. Sci. Technol. B*, **5**(1), 135–141.

Sasaki T (1979), *Conference on Very Large Scale Integration: Architecture, Design, and Fabrication*, California Institute of Technology.

Slodowski M, Doering H-J, Dorl W and Stolberg I (2011), Multi-shaped beam: development status and update on lithography results, *Proc. SPIE*, **7970**, 797013.

Stenkamp D, Kienzle O, Orchowski A, Rau W, Weickenmeier A, Benner G, Wetzke M, Waskiewicz W, Katsap V, Zhu X, Liu H, Munro E and Rouse J (2001), Progress on the realization of the electron column modules for SCALPEL high-throughput/alpha electron projection lithography tools, *Microelectron. Eng.*, **57–58**, 137–143.

Takizawa M, Komami H, Kurokawa M and Yamada A (2011), Position accuracy evaluation of multi-column e-beam exposure system, *Proc. SPIE*, **7970**, 79700B.

Trotel J (1978), Dynamic beam shaping, *J. Vac. Sci. Technol.*, **15**, 872–873.

van den Berg C, de Boer G, Boschker S, Hakkennes E, Holgate G, Hoving M, Jager R, Koning J, Kuiper V, Ma Y, van Mil I, Mook H, Ooms T, van de Peut T, Postma

S, Sanderse M, Scheffers P, Slot E, Tudorie A, Valkering A, Venema N, Vergeer N, Weirsma A, Woutersen S, Wieland M and Kampherbeek B (2011), Scanning exposures with a MAPPER multibeam system, *Proc. SPIE*, **7970**, 79700D (2011).

van Oven J, Berwald F, Berggren K, Kruit P and Hagen C (2011), Electron-beam-induced deposition of 3-nm-half-pitch patterns on bulk Si, *J. Vac. Sci. Technol. B*, **29**(6), 06F305–1–06F305–6.

Vistec Lithography, Ltd. (2012), http://www.vistec-semi.com/

Whitesides G, Mathias J and Seto C (1991), Molecular self-assembly and nanochemistry – a chemical strategy for the synthesis of nanostructures, *Science*, **254**(5036), 1312–1319.

Wieland M J, Kampherbeek B J, Addissi P and Kruit P (2001), Field emission photocathode array for multibeam electron lithography, *Microelectron. Eng.*, **57–58**, 155–161.

Winograd G, Meisburger W and Pease R (1999), Space-charge-induced aberrations, *J. Vac. Sci. Technol.B*, **17**(6), 2803–2807.

Yamada A, Yasuda H and Yamabe M (2008), Electron beams in individual column cells of multicolumn cell system, *J. Vac. Sci. Technol. B*, **26**, 2025–2031.

Yamada A, Oae Y, Okawa T, Takizawa M and Yamabe M (2010), Evaluation of throughput improvement by MCC and CP in multicolumn e-beam exposure system, *Proc. SPIE*, **7637**, 76370C (2010).

Yasin S, Hasko D and Ahmed H (2002), *Microelectron. Eng.*, **61–62**, 745–753.

4

Focused ion beams for nano-machining and imaging

M. UTLAUT, University of Portland, USA

DOI: 10.1533/9780857098757.116

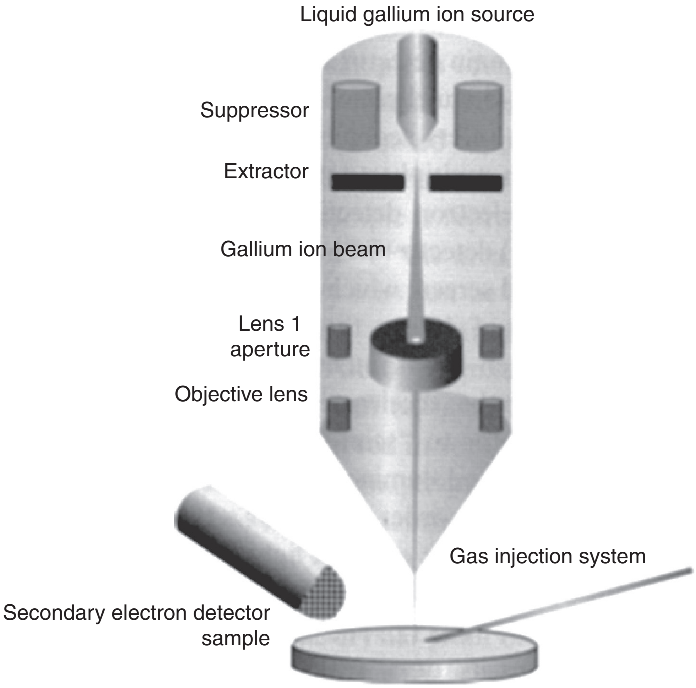
Abstract: This chapter discusses focused ion beams (FIB) and how they can be utilized for nano-machining and imaging. Following a brief history of FIB, four different types of ion sources are described, and the optics used to focus them. Following a description of the interactions of ions with matter that cause the removal of material by sputtering, which is useful for micro-machining, and the generation of electrons, which are useful for imaging, examples of ion milling, deposition of material, imaging, and spectroscopy at the nanoscale are given.

Key words: focused ion beam (FIB), milling, deposition, ion induced imaging.

4.1 Introduction

Nanotechnology is an expanding research and technological frontier where FIBs can be significantly utilized. FIBs produce tiny beams of a variety of positive ions (diameters from sub-nanometer to a few micrometers) energetic enough (from a few keV to ~ 150 keV) to cause the removal of sample material by elastic collisions (sputtering) at currents ranging from sub-pA to μA. This milling capability is localized on the sample at sizes from nanometers to microns and is at present the mainstay of work for FIBs. In addition to milling, ions can be selectively implanted into the sample. With the introduction of gases near the surface of the sample, material (either metals or insulators) can be deposited onto the sample at beam size dimensions to form novel structures. Due to the interaction of the ions with the sample through inelastic collisions, secondary electrons and ions are emitted from the sample. The FIB can be used as a microscope yielding information different from that of a scanning electron microscope (SEM), and by collecting and separating the secondary ions produced by the beam, secondary-ion mass spectrometry (SIMS) can be performed at high spatial resolution.

At present, Ga-based liquid metal ion sources (LMIS) are predominantly used, mainly due to the ease of their manufacturability, the stability of their



4.1 Schematic of the major components of a modern FIB system. The system consists of a focusing column containing the ion source and lenses, a detector, gas injection needles, and a sample stage, which are all inside a vacuum chamber.

operation, and the lack of need to heat the source, since Ga is a liquid at room temperature. The principal components of a FIB source and focusing optics are shown in schematic form in Fig. 4.1 for a typical Ga LMIS-based source with a two-lens focusing optical column equipped with a secondary-electron detector for sample imaging, and a gas delivery system for deposition. All that is shown is contained in a vacuum chamber.

By inserting a Wien filter mass separator into the optical column, alloy sources can be made to operate so that a variety of ions (e.g., B, Be, Li, Si, Au) are selectively available to form the beam.

With the recent commercial introduction of gas field ionization sources (GFIS), He and Ne are available as ions, and have shown considerable promise for use in producing very high resolution images and for high precision nano-machining. The He-based microscope has demonstrated 0.35 nm imaging resolution, with contrast mechanisms unavailable to the SEM, while Ne based systems are demonstrating very precise nano-machining capability.

For the removal of a large volume of sample, higher beam currents are necessary to carry out the work in a reasonable time. Due to limitations of the LMIS sources, plasma-based sources have been introduced that are capable of producing beam currents of up to a few μA. These systems (now available from several commercial sources) excite plasmas from which ions are extracted and focused, extending the capability of FIBs to work in relatively high volume sample removal regimes unavailable to the LMIS- and GFIS-based systems.

To the future of FIB work belongs the cold-atom ion sources (CAIS), which are magneto-optical trap (MOT)-based ion sources that should be capable of producing a variety of different ions, with some early demonstrations and calculations indicating that sub-nanometer imaging resolution should be achievable. In addition, by switching ion types in the trap, it is expected that milling capabilities should also be achievable at the nanometer scale.

A major technological success has been to marry the FIB with SEM in a 'dual beam' system. Thus, in one chamber, all the advantages of each technique are combined. The value of this blended system is evident in the number of sales (~thousand) that manufacturers have achieved.

Following a brief history of past efforts, we describe the ion sources and the optical columns used to focus and manipulate the beam. After detailing the interactions of the ions with the sample, we show examples of milling, deposition, imaging, and spectroscopy.

4.2 An adumbrated history of focused ion beams (FIBs)

The short history that follows is by necessity inadequate, due to the omission of many workers deserving mention. Ga-based FIB is a mature technology, having roots in scanning electron beam, probes ion thrusters, and ion implanters. Following work on SEMs, ion systems capable of producing 0.1 μm ion beam probes of a few picoamperes from low-brightness plasma sources were developed (Hill, 1968). Crewe (Crewe *et al.*, 1968; Crewe, 1973), showed the importance of high-brightness sources and low-aberration optical components for high resolution systems. Systems based on the GFISs were developed at Chicago and Oregon (Levi-Setti, 1974; Orloff and Swanson, 1975). Due to technological limitations, this approach was temporarily abandoned in favor of the LMIS.

The first LMIS-based focusing columns built in 1975 (Krohn and Ringo, 1975) were followed by work at Hughes Research Laboratories (HRL) (Seliger *et al.*, 1979), where several versions of the early FIB systems were intended to be used for the maskless implantation of semiconductors. Many groups then followed those efforts in exploring the use of FIBs in the areas of microscopy, spectroscopy, lithography, mask repair, TEM sample preparation, and microcircuit surgery. What began as an implantation tool has found its main role in the sputter removal of material and, to a lesser extent, for the deposition of material and as an imaging tool, with the Ga LMIS as the workhorse source type.

At present, Ga-based FIB is a mature technology and is used mainly for the preparation of TEM samples, failure analysis for semiconductors and MEMS, and circuit edit procedures, as well as a tool for MEMS and nanotechnology proto-typing applications. Some investigations are still in the process of determining whether ions different from Ga- and LMIS-based sources will find widespread applications.

Recent efforts (Notte and Ward, 2006) in producing a working He ion GFIS have pushed the frontiers of ion microscopy to the limit, showing spectacular imaging capabilities. More recent work (Tan *et al.*, 2011) on GFIS has produced a system, currently under exploration, capable of delivering near-nm sized beams of Ne onto a sample for precision milling applications.

For applications requiring very high currents or non-Gallium ions, a new FIB system based on an inductively coupled plasma ion source, producing Xe and O ions at a beam current of a few μA, has been developed (Smith *et al.*, 2006; Kellogg *et al.*, 2010).

Recent attempts to extract ions from MOTs have shown promise that these sources, which produce ultra-cold atoms that can then be subsequently ionized and focused, may be useful for ultra-high resolution imaging and milling (Reijnders *et al.*, 2009; McClellen *et al*, 2011). At present, this is a future technology whose value has yet to be proven.

4.3 Sources of ions: a quartet of types

The heart of the FIB is the source of ions, and there are four types of ion sources that we will discuss here: (1) the LMIS, (2) the GFIS, (3) the inductively coupled plasma FIB source (PFIB), and (4) the CAIS based on the MOT. The most used measures of the quality of an ion beam are its 'spot' size and current density. The quality of the ion beam obtained depends not only on the properties of the source, but also on the properties of the optical system (the 'optical column,' or 'column'), and also the Coulomb interactions of the ions in the beam (Ward *et al.*, 1987, 1988, 1991). In most FIB systems, the optical column de-magnifies the source and images it on the sample. For example, a source that is 50 nm in size will, upon a demagnification of 20, be 2.5 nm in size at the sample, in the absence of other aberrations. From an optical perspective, the most important parameters to consider for the ion sources are their brightness, angular intensity, and energy spread.

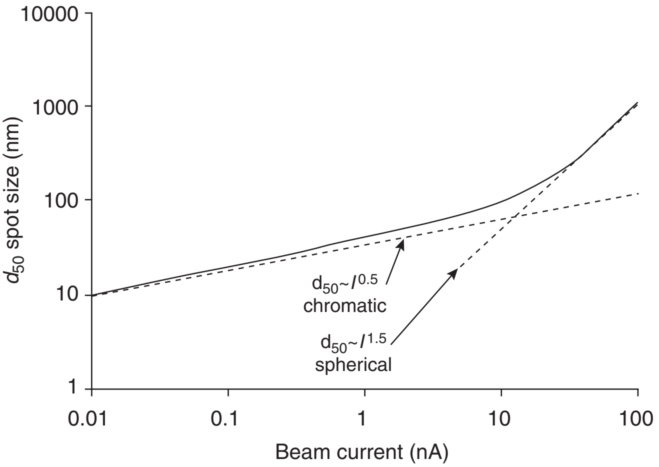
The 'reduced' or 'specific' brightness B_r (A/m^2 × S_r × V) of sources (Tandare, 2005) is defined as

$$B_r = \frac{I}{\delta \Omega \mathrm{V}} \quad [4.1]$$

where I is the source ion current, δ the virtual source size, Ω the solid angle into which the ions are emitted, and V is the acceleration voltage. Reduced brightness is conserved throughout the optical system, so that its value at any focus of the ion beam stays the same irrespective of the imaging conditions. The optical properties of the column (which transfers and de-magnifies an image of the source of ions to the sample) are determined by the aberrations of the optical elements, which typically comprise electrostatic lenses. There are four main aberrations that must be considered: chromatic, spherical, diffraction, and finite source size. Diffraction is not an issue for FIB, since the wavelength of typical ions used is so small: for example, a 30 keV Ga^+ ion has a wavelength of 26.5 fm, about five orders of magnitude less than the other aberrations. The effect of chromatic aberration in optical systems is for lenses to bring ions of different energy to focus at different points along the optical axis, with higher energy ions being brought to focus further from the lens than those of lower energy. If the energy spread of the ions is large, the lens has no unique focal length and gives rise to a 'disk of confusion' rather than a point focus, deleteriously affecting the resolution attainable with the optical system. Hence, a small value for the energy spread of the source will give a smaller focused beam at the sample. In the absence of energy spread, spherical aberration limits the ability of the optical system to form a point focus for a large diameter 'monochromatic' ion beam. Ions farther from the optical axis are brought to focus sooner than those near the axis. In order to minimize the effect of spherical aberration in an optical system, only the rays that are near the optical axis are used (i.e., an aperture is inserted into the beam path). Since the size of the aperture will determine how many ions pass through it, there is a sesquilinear relationship between the current in the beam and the amount of spherical aberration. This relationship is manifested in the growth of the beam size with increasing beam current; more current means a larger diameter beam. A larger angular intensity from the source allows more current into the final beam at a given aperture size. These effects can be seen in Fig. 4.2, where the two regimes, chromatic and spherical, are evident.

4.3.1 Liquid metal ion sources (LMIS)

By far, the Ga LMIS is the most extensively used of the four source types (Orloff *et al.*, 2003; Giannuzzi and Stevie, 2005; Yao, 2007). The LMIS is in a class known as electrohydrodynamic sources (Swanson, 1994) and for FIB



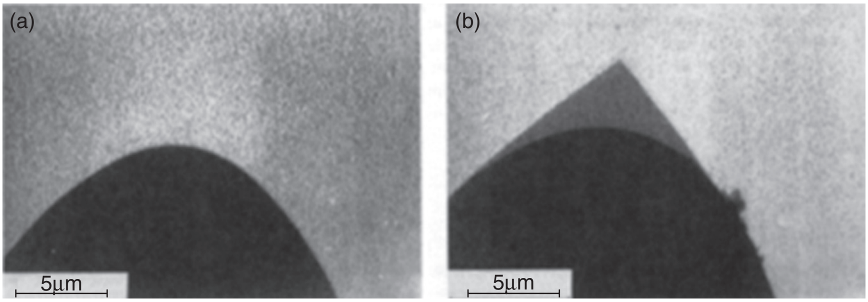
4.2 The d_{50} spot size is the diameter of the beam that contains 50% of its current, I. Shown is the d_{50} spot size vs the beam current. Since the beam current is determined by the diameter of a defining aperture d_A, the growth of the spot size with current from 0.001 nA to about 10 nA is dominated by chromatic aberration (scales as d_A), and above that by spherical aberration (scales as d_A^3).

they are realized when an electrochemically etched, blunted field emitter needle (usually W) with an end radius of ~10 μm coated with a liquid metal heated to its melting point is placed in an electric field produced by a proximally placed extraction electrode at negative voltage (~5–15 kV) relative to the needle. The liquid metal is pulled into a conical shape (see Fig. 4.3), the 'Taylor cone,' by the balance of electrostatic and surface tension forces.

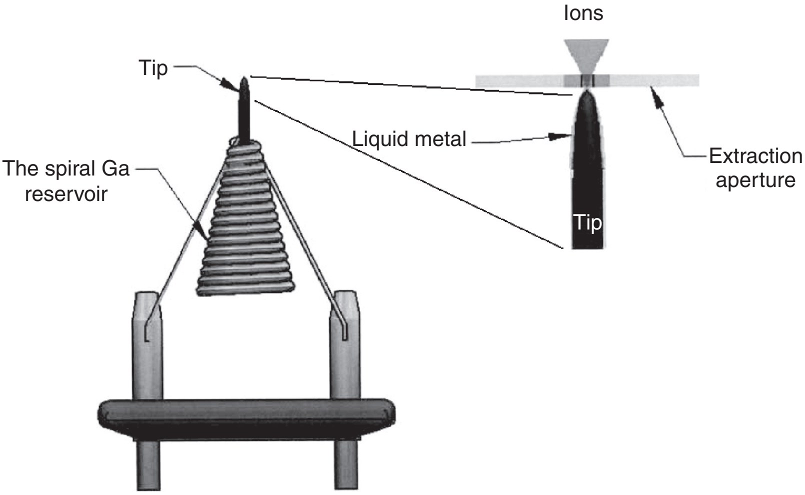
The end radius of the liquid apex is small enough that the electric field at the apex causes ions to form through the mechanism of field evaporation (Kingham and Swanson, 1984).

Field evaporation is the dominant mechanism of emission at low currents (few μA), where LMISs should be operated in FIBs, and occurs when either an atom or ion is evaporated from the surface of a metal through a potential barrier that has been lowered by an external electric field. If an atom has been evaporated, it becomes an ion by post-ionization. Figure 4.4 shows a schematic of such a source.

The main advantages of the Ga LMIS are: (1) Ga is a liquid at room temperature, so the source does not require heating in order to operate; (2) Ga readily wets tungsten, which is the preferred needle substrate; (3) the brightness, energy spread, and angular intensity of the Ga source are all reasonable values with which optical columns can be made to transport the ions to the sample (Orloff *et al.*, 2003).



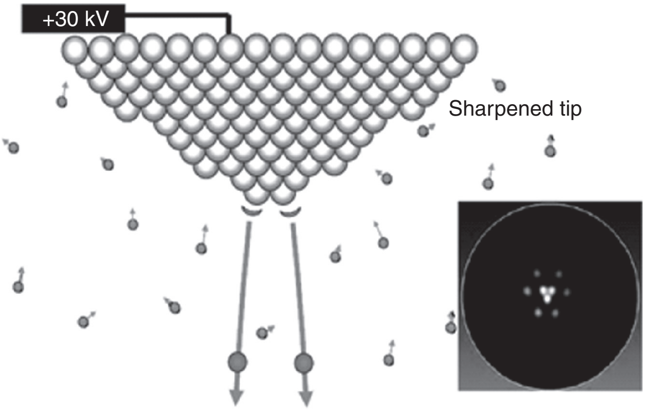
4.3 Transmission electron micrographs (Driesel *et al.*, 1996) of a AuGe LMIS emitter. The image (a) is of the W emitter substrate. The image (b) is the emitter substrate with the Taylor cone.



4.4 Schematic of a typical LMIS source showing the needle, reservoir, heater, and base. The base contact pins allow current to flow through the reservoir to melt the metal which then flows down the needle from which ions are then extracted by the electric field created by the extraction electrode.

4.3.2 Gas field ionization sources

The GFIS ionizes gas in the high electric field near a sharpened metal tip (Notte and Ward, 2006). A gas based, *in situ* process strips individual atoms from a sharpened needle (usually tungsten) until an atomic sized pyramid is created with three atoms at the end of the source tip. Once this 'trimer' is formed, the tip is operated in high vacuum at cryogenic temperature with helium gas flowing over it. High voltage applied to the tip produces a very



4.5 Schematic of the He GFIS source. The inset shows an image of the 'trimer' of three atoms formed at the apex at were the He atoms become ionized.

high electric field at the apex, which attracts the helium gas to the energized tip where it is ionized, as shown schematically in Fig. 4.5.

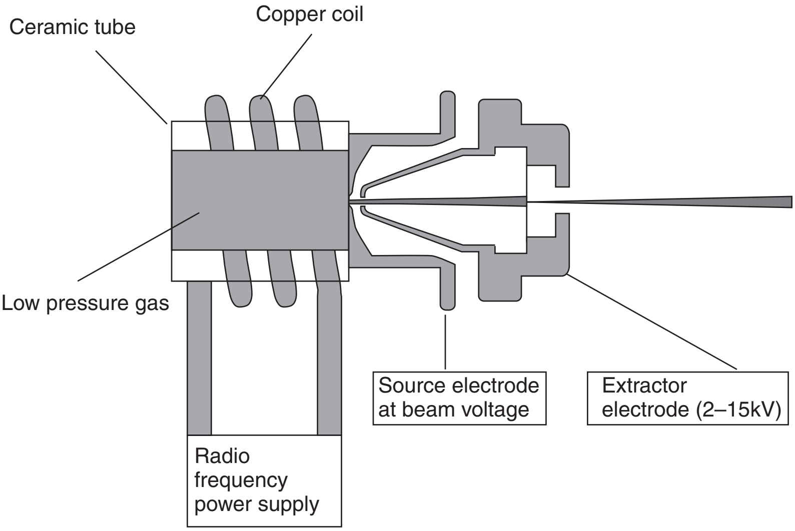
Field ionization happens in the vicinity of a single atom at the apex, so that the resulting ions appear to be emanating from a region that is less than an angstrom in size. This produces a high-brightness beam of ions, which are extracted and focused with an optical column to sub-nanometer probe size. Recently, in addition to helium, GFIS has been made to produce neon ions in a similar manner (Tan *et al*, 2011).

4.3.3 Inductively coupled plasma ion source (PFIB)

There are several ways to produce a source based on creating a plasma from which ions are extracted, and we discuss here one such method. The inductively coupled plasma ion source utilizes an 'antenna' operating at RF frequencies to cause plasma to form in a small chamber by Faraday induction (Smith *et al*, 2006; Kellogg *et al.*, 2010). A noble gas at low pressure in the chamber is excited into the plasma state by the cylindrical RF antenna, from which ions are extracted by means of a biased electrode as shown in Fig. 4.6.

4.3.4 Cold atom ion source (CAIS)

The CAIS is based on an MOT and is expected to have a significantly smaller energy spread than the LMIS at comparable brightness, enabling <1 nm focused spot sizes. Additionally, the CAIS has the potential to

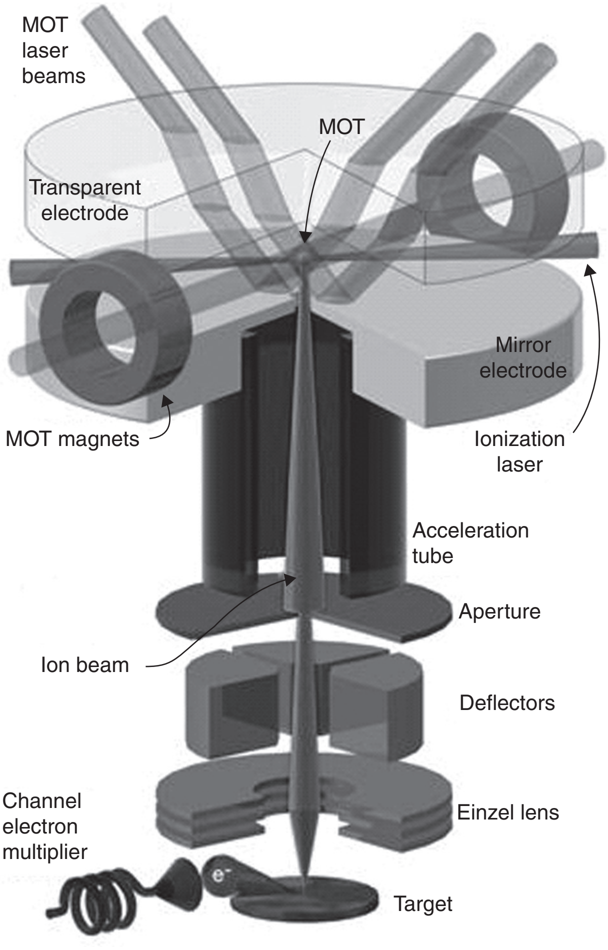

Ceramic tube
Copper coil
Low pressure gas
Source electrode at beam voltage
Extractor electrode (2–15kV)
Radio frequency power supply

4.6 Schematic of an inductively coupled ion source. A gas filled ceramic tube surrounded by an RF 'antenna' which when energized causes a plasma to form in the tube. A series of electrodes suitably biased extract ions from the plasma. Typically, Xe, Ar and O gases are used. While there are other types of 'plasma' source configurations, this one offers superior brightness, angular intensity and virtual source size . This type of source has a much larger source size than the other three sources considered here (see Table 4.1). Successful sources of Xe and O have been made.

produce many different kinds of ions, as compared to the other ion sources discussed, in particular the alkali metals such as Cs for use in time-of-flight SIMS with both high spatial resolution and high sputter yield, and Li for imaging purposes as a future alternative to He in the GFIS-based He ion microscope.

This type of source is a laser-cooled and trapped atomic gas inside an MOT, which typically consists of a vapor cell where the radiation pressure exerted by three orthogonal pairs of laser beams of pair-wise opposite circular polarizations within a quadrupole magnetic field cool the atoms to temperatures below 1 mK, causing the atoms to collect in a cloud suspended in the region at the center of the trap where the magnetic field vanishes. The gas is ionized by near-threshold photo-ionization and, upon applying an extraction electric field to the trapped atoms, a cold ion beam is created and extracted, as shown schematically in Fig. 4.7.

The thermal energy of these ions (100 μK ~ 10^{-8} eV) is very low, giving rise to a very high brightness, mainly due to the small solid angle into which the ions are emitted (Reijnders *et al.*, 2009; McClellen *et al.*, 2011).


MOT laser beams
MOT
Transparent electrode
Mirror electrode
MOT magnets
Ionization laser
Acceleration tube
Aperture
Ion beam
Deflectors
Channel electron multiplier
Einzel lens
e⁻
Target

4.7 Schematic of a MOT-based CAIS system. The source has a MOT with an extractor which feeds an accelerator region and a focusing lens. Typically this type of source will produce alkali metals such as Li for imaging, and Cs and Rb for milling. The thermal energy of these ions (100 μK ~ 10^{-8} eV) is very low, giving rise to a very high brightness mainly due to the small solid angle into which the ions are emitted (Reijnders *et al.*, 2009; McClellen *et al*, 2011).

4.4 Charged particle optics

Once a beam of ions has been produced in a source, it must be transported to the sample by an optical column. The purpose of the optical column is multifold: source demagnification, ion acceleration, focusing, and scanning. Most FIB optical columns employ two electrostatic lenses in order to be able to de-magnify the source, focus on the sample, change beam energy,

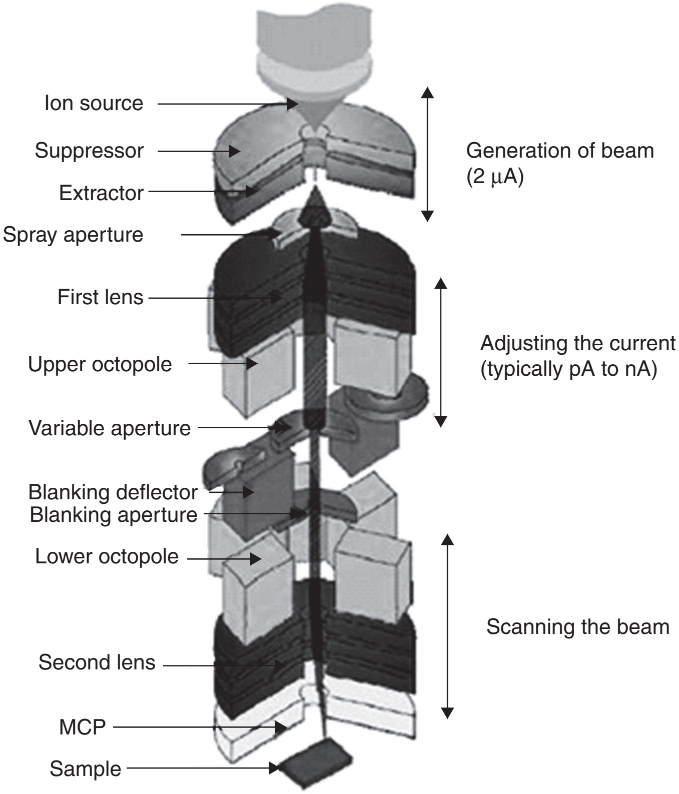
Table 4.1 The four kinds of ion sources, and some of their parameters

	Presently used	Presently used	Future use	Future use
Source type	LMIS	GFIS	PFIB	CAIS
Ions available	Ga, In, Au, Si, B, Be, As	He, Ne	Xe, O	Li, Rb, Cs
Ion energy (keV)	1–50	2–50	2–50	1–50
Reduced brightness A/(m^2 srV)	1×10^6	1×10^7	9×10^3	5×10^5
Source size (nm)	50	0.2	6000	20
Source energy spread (eV)	5 for Ga	<1	7 for Xe	1
Source angular intensity (mA/sr)	0.02		7.4	?
Ion beam current	1 pA–100 nA	1 pA–few pA	~10 pA–few μA	Few pA?
Primary Uses	Mill, depo, image	Image, mill, depo	mill	Image, mill

Note: LMIS. GFIS. Plasma FIB source (PFIB). CAIS.

and vary the beam current, as shown schematically for a Ga-based LMIS system in Fig. 4.8.

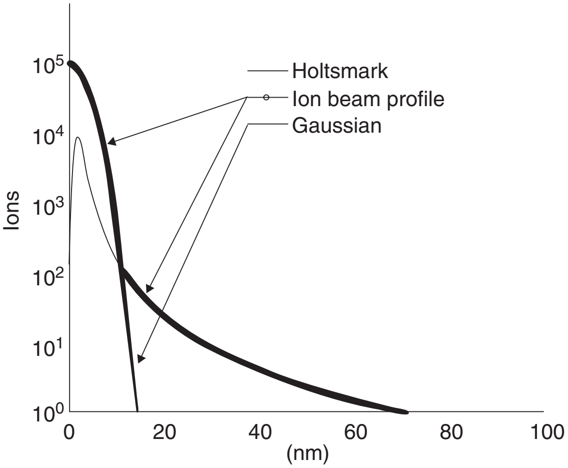
There are well-known design principles to reduce lens aberrations (Grivet, 1972) and computer programs to simulate optics (Munro, 1973; Dahl *et al.*, 1990). However, the problem of determining the effects of Coulomb interactions between charged particles in the beam is very complicated. The Coulomb interactions depend on several system parameters, including ion mass, beam energy, beam current, beam convergence angle, source properties, length of optical column, crossover size, and position (Jaing and Kruit, 1995). The effects caused by Coulomb interactions are of three types: space charge, Boersch effect, and trajectory displacement. Space charge is a spatial broadening of the beam caused by its average charge density, and acts like a defocusing negative lens. This defocusing can mostly be corrected in the optical system. The Boersch and the trajectory displacement effects lead to axial and lateral displacements of the charged particles, respectively. The Boersch effect results in energy broadening, and the trajectory displacement effect results in broadening the spatial ion distribution. Both of these effects are statistical in nature, due to the Coulomb repulsion of individual pairs of the ions. The deleterious results of the statistical effects cannot be corrected in the optical system, and can only be minimized by proper design. Low energy high current density FIBs are desirable for low damage maskless processing, but Coulomb interactions


Ion source
Suppressor
Extractor
Spray aperture
First lens
Upper octopole
Variable aperture
Blanking deflector
Blanking aperture
Lower octopole
Second lens
MCP
Sample
Generation of beam
(2 μA)
Adjusting the current
(typically pA to nA)
Scanning the beam

4.8 Cut-away schematic of a generic two-lens optical column used with LMIS-based sources. About 2 μA of source current is accelerated and reduced by an aperture early along the beam path to limit Coulomb interactions and then is received by a condenser lens and variable current defining aperture which sets the final beam current. The upper octupole is used to 'blank' the beam. The final section of the column focuses and scans the beam (lower octupole) on the sample.

can dominate in this regime. The energy spread in FIBs can reach widths of several eV, even if the initial energy spread at the source is only a fraction of an eV. The net result of Coulomb broadening is to alter the beam profile from a Gaussian-like shape to one that is more Holtsmark-like (Chandrasekhar, 1943; Ward *et al*., 1987, 1988). Measured beam profiles have the shapes shown in Fig. 4.9.

For sources that emit different species of ions, a mass filter must be inserted into the beam path in order to select a particular ion (Seliger, 1972). Many LMISs are alloy- or compound-based, and require an optical column such as shown schematically in Fig. 4.10. For a more comprehensive review of such systems, see (Orloff *et al*., 2003).



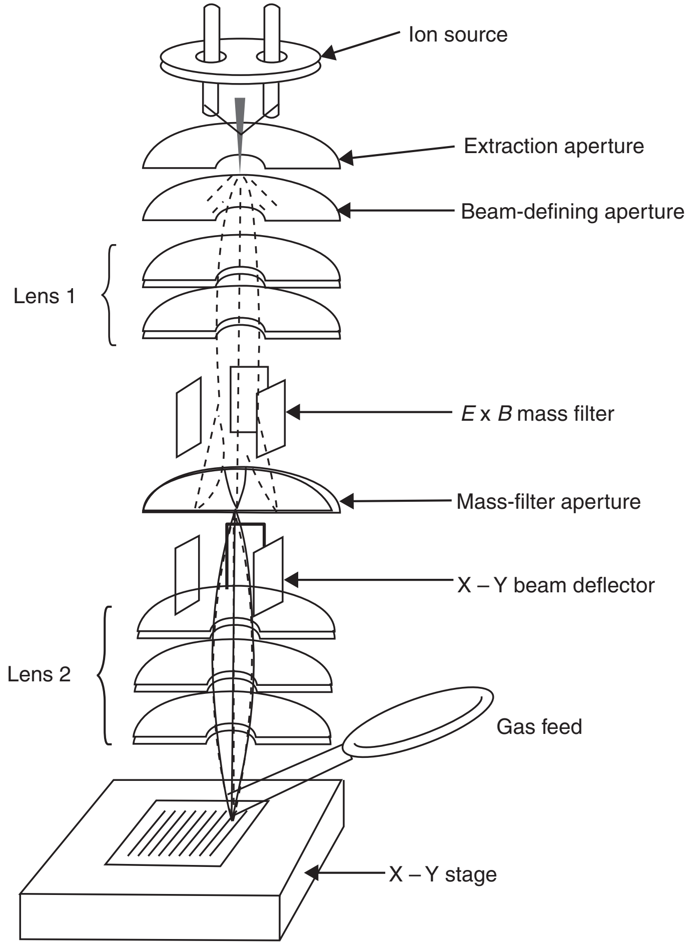
4.9 The shape of the ion beam is important for milling, deposition and imaging. Due to Coulomb interactions the beam is not Gaussian in shape. The central part of the profile is Gaussian-like, but the Holtsmark character (Ward *et al.*, 1987, 1988) of the profile leads to 'tails' which in some applications become important.

4.5 Ion-matter interactions

After extracting ions from a source and focusing and conveying them to the sample, it is important to understand the events that impinging ions are capable of initiating, and we discuss here their interactions with matter. Fig. 4.11 shows a schematic of the various generic ion-matter interactions. The basic interactions are that sample atoms may be removed by sputtering (Sigmund, 1969, 1988; Benninghoven *et al.*, 1987), secondary electrons may be produced, damage may be incurred by amorphization or mixing, and gas chemistry may be initiated to cause deposition onto the sample or cause selective and accelerated milling.

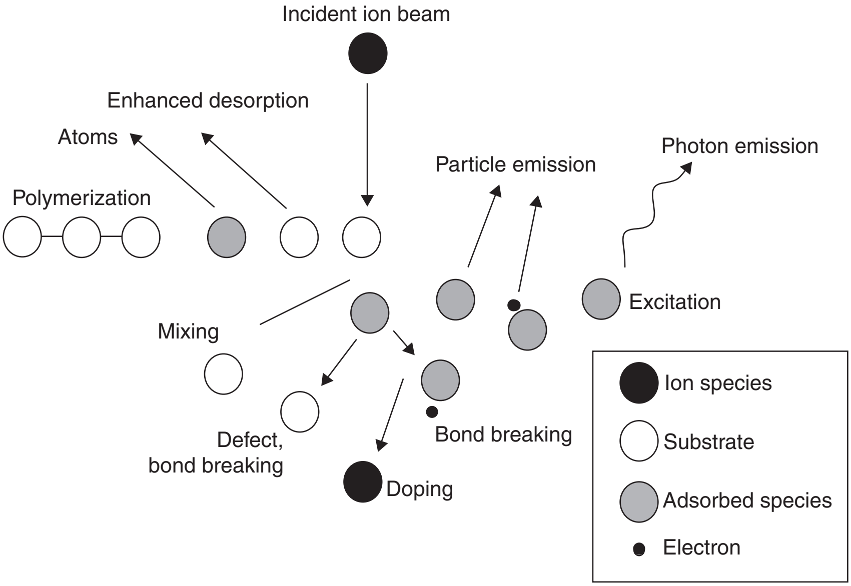
As primary ions in the impinging beam enter the sample, they lose energy through two channels (nuclear and electronic), and come to rest after all of their initial energy is lost. The path of such a typical ion is shown in Fig. 4.12, where the range R is the total path length the ion takes as it loses all of its initial energy (and is usually longer than the projected range), and R_p is the final resting depth of the ion. The transverse straggle, R_t, is the sideways projection of the range.

Nuclear losses are due to collisions between the screened nuclear charges of the incident ion and target atoms, whereas electronic losses are due to incident ions interacting with electrons in the sample resulting in excitation and ionization. The nuclear channel causes physical sputtering, which modifies the structural landscape of the sample by material removal and

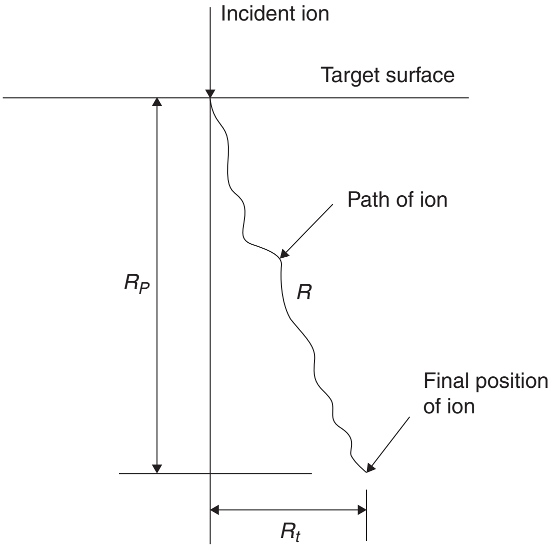

Ion source
Extraction aperture
Beam-defining aperture
Lens 1
E x *B* mass filter
Mass-filter aperture
X – Y beam deflector
Lens 2
Gas feed
X – Y stage

4.10 Schematic of a two-lens mass separated column. An $E \times B$ (Wien) filter is inserted into the beam path between the lenses. The mass filter passes one velocity (corresponding to one mass) through an aperture and deflects other masses so that they are blocked by the aperture. The mass filter is necessary when alloy or compound sources are used.

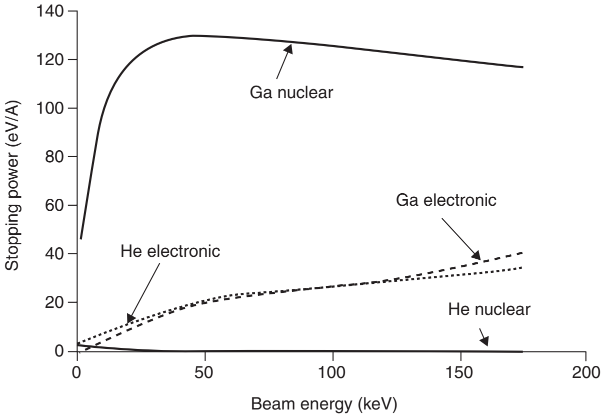
results when the incident ions transfer sufficient momentum to surface or near-surface atoms to free one or more target atoms. The sputtering process can be chemically modulated with the addition of gases into the system, whereby chemical sputtering is initiated for increased or selective removal of sample material. Concomitant with material removal is generation of sample information through the production of secondary ions and electrons from the sample by electronic losses. Since these secondary quanta are fairly localized to where the ion beam strikes the sample, they can be used to form images of the sample with a resolution of nearly the beam


Incident ion beam
Enhanced desorption
Atoms
Polymerization
Particle emission
Photon emission
Excitation
Mixing
Defect, bond breaking
Bond breaking
Doping
Ion species
Substrate
Adsorbed species
Electron

4.11 The events that an incoming ion can cause are: implantation, sputtering, generation of secondary electrons and ions, and the induction of gas-sample interactions.


Incident ion
Target surface
Path of ion
R_P
R
Final position of ion
R_t

4.12 The geometry of the range R, the projected range R_p, and the transverse straggling distance R_t for an ion penetrating into the sample. A rule of thumb for heavy ions (Ar, Ga, Xe) is that $R \sim 1$ nm/keV.



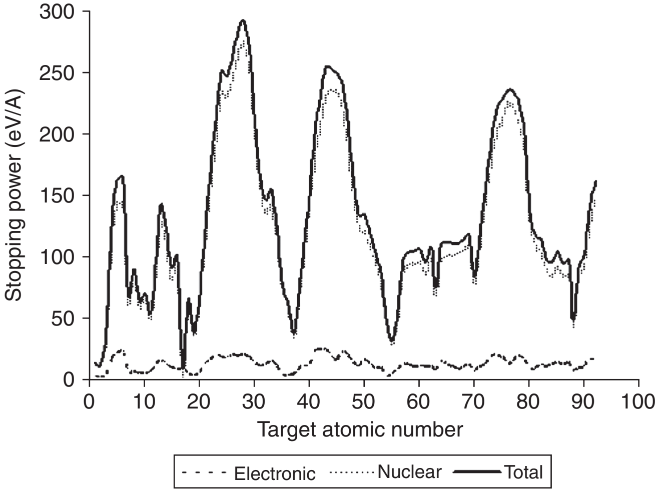
4.13 The nuclear and electronic stopping powers for He and Ga impinging into Si as a function of energy. The virtual lack of nuclear loss for He and a large one for Ga results in almost no sputtering for He, and significant sputtering for Ga. The electronic losses for both He and Ga are inelastic, and generate secondary electrons. Because the losses for He are much less than those for Ga, He has a much larger range.

size. Secondary-ion emission occurs when surface atoms are ionized, and sufficiently energized to be emitted from the surface. The electronic channel causes secondary electrons to be emitted when the interaction is near the surface, and happens mostly from ion bombardment of metallic surfaces, with the typical yield per incident ion 10–1000 times greater than secondary-ion yield. The two types of signal yield different information about the sample, and are also different from SEM images (see Section 4.7).

The total stopping power for the ions is the sum of the nuclear and electronic losses. It depends on the mass of the impinging ion and the mass of the target atoms, as well as the energy of the incoming ion. Generally, heavier ions (e.g., Ga, Xe) lose most of their energy via the nuclear channel, while lighter ions (e.g., He, Li) tend to lose energy predominantly through the electronic channel. Hence, for most materials of interest, Ga and He lose energy differently, and that is shown for the particular case of Si in Fig. 4.13, where it should be noted that the absolute loss rate is much higher for Ga than He, and that is why He penetrates much further into the sample, and also why Ga has a much larger sputter yield.

It can be shown that a useful relationship for calculating the sputter yield is given by

$$Y(E_0) = \frac{4.2\times 10^{14}}{U_s}\alpha S_n \qquad [4.2]$$



4.14 The electronic, nuclear and total stopping power of 30 keV Ga$^+$ across the periodic table. There is a correlation of the nuclear loss with both sputtering and secondary-electron yield.

where Y is the sputter yield, E_0 the energy of the incident ion, S_n the nuclear stopping power, and U_s the sublimation energy, and

$$\alpha = 0.15 + 0.13 \frac{M_2}{M_1} \qquad [4.3]$$

where M_1 and M_2 are the masses of the incident ion and the target atom. Furthermore,

$$S_n = \frac{8.5\, s_n(\varepsilon) M_1}{M_1 + M_2} \frac{Z_1 Z_2}{\sqrt{Z_1^{2/3} + Z_2^{2/3}}} \qquad [4.4]$$

where Z_1 and Z_2 are the atomic numbers of the incident ion and the target atom, and $s_n(\varepsilon)$

is given by

$$s_n(\varepsilon) = \frac{0.5 \ln(1+\varepsilon)}{\varepsilon + 0.14\, \varepsilon^{0.42}} \qquad [4.5]$$

with ε the 'reduced energy' given by

Table 4.2 The density, and for 30 keV Ga^+ ions the sputter rate and sputter yield are shown for a variety of sample materials of interest

Element	Density (g/cm^3)	Sputter rate (μm^3/nC)	Sputter yield (atoms/ion)
C (diamond)	3.57	0.18	2.7
Al	2.7	0.3	2.9
Si	2.33	0.27	2.1
Ti	4.5	0.37	3.3
Cr	7.19	0.09	1.2
Zn	7.13	0.34	3.6
Ge	5.32	0.22	1.6
Se	4.81	0.43	2.5
Mo	10.2	0.12	1.3
Ag	10.5	0.42	3.9
Sn	5.76	0.25	1.2
W	19.25	0.12	1.2
Pt	21.47	0.23	2.4

$$\varepsilon = \frac{32 M_2 E\ [keV]}{Z_1 Z_2 (M_1 + M_2)\sqrt{\left(Z_1^{2/3} + Z_2^{2/3}\right)}} \qquad [4.6]$$

Once the sputter yields Y are known, the rate of sputtering Y_r (volume removed/incident charge) can be calculated from

$$Y_r\left(\frac{\mu m^3}{\mathrm{nC}}\right) = \frac{\mathrm{MY}}{96.4\rho} \qquad [4.7]$$

where ρ is the sample density in g/cm^3 (Benninghoven *et al.*, 1987).

This is a useful relationship for estimating the time for sample volume removal. For example, suppose one wants to remove 1 μm^3 of Si with a 30 keV Ga^+ beam. Since the charge q = current μ time and the sputter yield is 2.1 atoms/ion, giving a sputter rate of 0.28 μm^3/nC, one needs 3.57 nC delivered to the sample. For a beam current of 1 pA the time required would be about 3600 s (an hour) while a 10 nA beam would accomplish the same task in 0.36 s. The time necessary to mill is not the whole story (Orloff *et al.*, 2003; Yao, 2007). One must also consider scan strategies to reduce re-deposition and, in the example just given, the beam profiles would be quite different, with the larger current beam being significantly larger in size. Table 4.2 gives sputter yields and rates for some materials when bombarded by 30 keV Ga^+ ions.

Calculating these ion–solid interactions can be difficult, but there exists a good computer simulation with which to estimate the range and the

Table 4.3 Tabulation of TRIM calculations for Ga^+ incident on Si at normal incidence for a variety of energies

keV	Vacancy/Ion	R_p (nm)	ΔR_p (nm)	ΔR_t (nm)	S (atom/ion)	eV/atom
50	1177	41.1	14.8	11.6	2.3	80
40	967	34.9	12.6	9.6	2.4	64
30	748	28.0	9.9	6.3	2.5	52
20	516	20.3	7.4	4.6	2.4	41
10	274	12.9	4.6	3.8	1.9	35
5	144	8.3	3.0	2.6	1.6	23
2	61	4.9	1.7	1.6	1.1	17
1	32	3.5	1.2	1.2	0.6	10

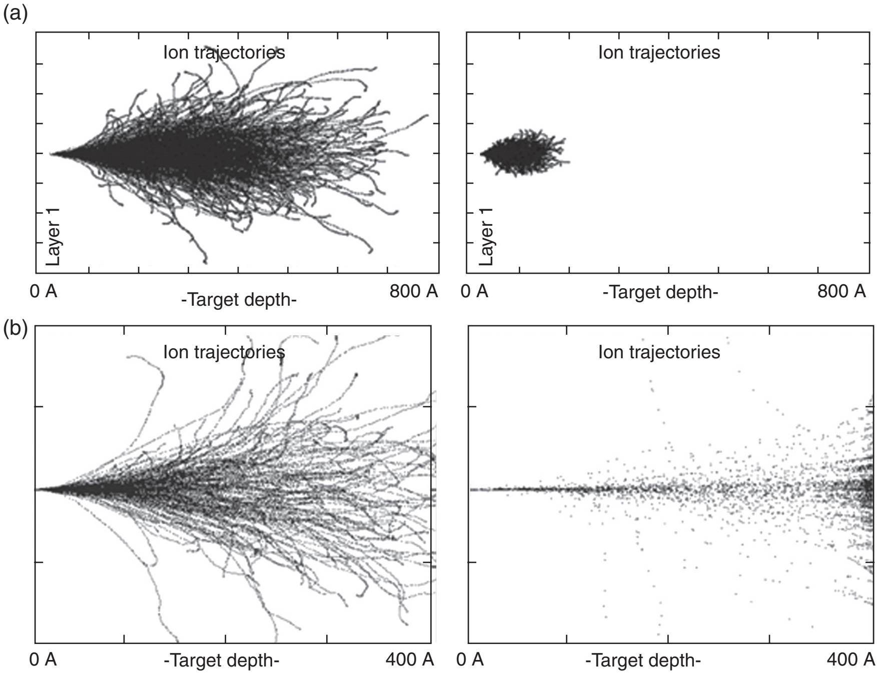
Table 4.4 Comparison of ranges, straggles, sputter yields and energy transferred to Si target atoms for several different incident ions at 30 keV normal incidence

30 keV	Range	Longitudinal straggle	Lateral range	Lateral straggle	Sputter yield	eV/atom
He	280	79	82	102	0.03	112
Be	158	58	57	70	0.11	25
Al	53.1	23.1	15.8	20.2	1.3	144
Ga	28.1	10.1	6.4	8.3	2.6	39
In	24.8	7.3	5.1	6.4	3.6	39
Xe	24	6.5	4.2	5.3	3.4	30
Bi	22.8	4.5	3.5	4.4	3.8	37

distribution of the extent to which ions have straggled both longitudinally and transversely from the mean range.

TRansport of Ions in Matter (TRIM) (Ziegler, 1991) is a computer program using a binary collision-based Monte Carlo simulation that calculates the three-dimensional penetration of ions into matter. These simulations are fairly accurate for estimates of the sputter yield (target atoms sputtered per incident ion) and damage incurred by the sample as ions impinge onto it. The program is useful for engineering solutions for specific applications, as both the sputter yield and damage are functions of incident ion species, energy, and angle. Such effects are shown in Figs 4.15a and 4.15b, in which the incident ion type and energy are changed. Table 4.3 shows a compilation of TRIM results for Ga^+ at normal incidence into Si at different energies. Table 4.4 shows a comparison of different 30 keV ions impinging onto Si.

Most calculations and simulations of ion–solid interactions assume that the sample is amorphous. In crystalline materials, a phenomenon known as channeling is possible, and occurs when the incident ions have correlated collisions with target atoms as the ions enter in a direction along a

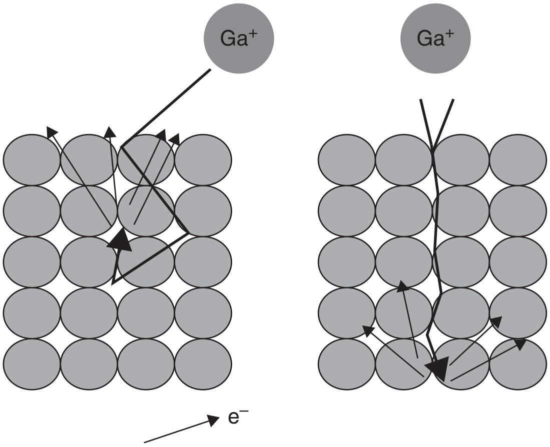


4.15 (a) TRIM calculations, showing the effect of incident energy. The left plot is of 40 keV Ga^+ ions into Si. The right plot is for 5 keV Ga^+ ions into Si. (b) TRIM simulation showing the effect of the impinging mass of the ion. The left plot is 30 keV Ga^+ into Si. The right plot is for 30 keV Be^+.

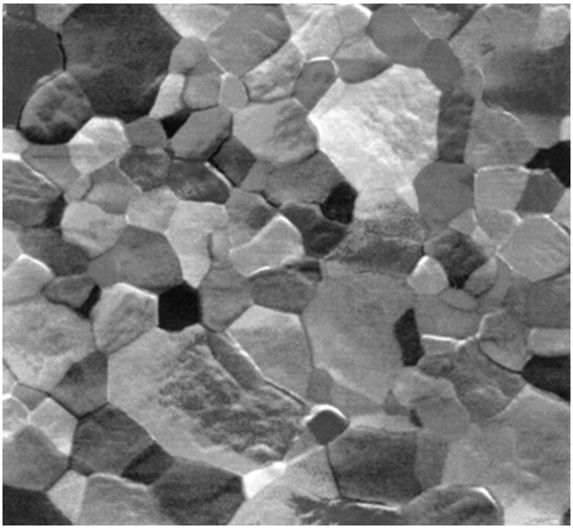
crystal axis. Steering of the ions due to open channels can result in ion ranges several times the maximum range in non-steering directions or in amorphous materials. When this occurs, electronic losses determine the range, and there is little straggling (hence the term ‘channeling’), and then near the end of their paths nuclear collisions with lattice atoms dominate the energy loss to stop the ions. This effect is illustrated schematically in Fig. 4.16, and an example of 30 keV Ga^+ impinging an Al sample in Fig. 4.17 where the channeling effect is obvious in the different orientations of the microcrystals.

The incoming ions produce damage to the sample, and in some cases the damage must be minimized. The most expedient route to lower the damage is to reduce the energy of the ions. Figure 4.18 shows the extent of the damage layer with varying energy in Si. It is apparent that the thickness of the damage layer can be suitably controlled by lowering the energy of the incoming ions.

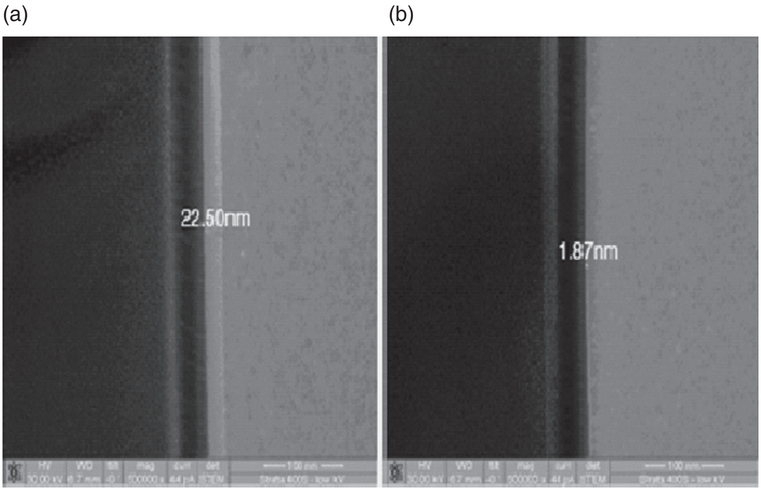
Enhanced milling (Fig. 4.19) is achieved when there is adsorption of the gas onto the substrate followed by an interaction of the gas molecules with the



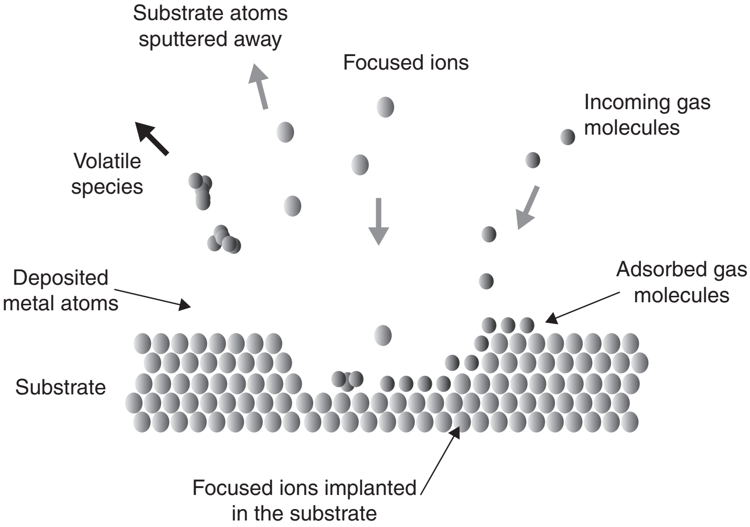
4.16 Illustration of the channeling effect. Ions not entering a crystal axis generate losses near the surface (generating electrons), while ions that 'channel down' a crystal axis lose energy deeper within the crystal with the concomitant loss of sputtering and secondary-electron generation near the surface.



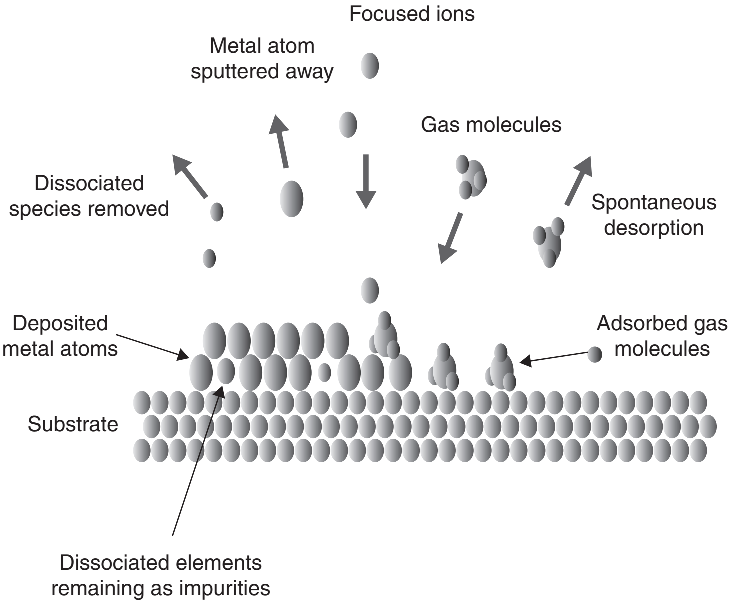
4.17 Channeling effect in Al by Ga. The differences in contrast are due to the different depths of penetration into the different indexed axes of the microcrystals in the field of view. Darker contrast means deeper penetration.


(a)
(b)
22.50nm
1.87nm

4.18 TEM micrographs of the depth of the damage layer for Ga⁺ into Si. The rule of thumb that the range is about a nanometer per keV of energy for Ga⁺ is seen here for (a) 25 keV and (b) 2 keV.


Substrate atoms sputtered away
Focused ions
Incoming gas molecules
Volatile species
Deposited metal atoms
Adsorbed gas molecules
Substrate
Focused ions implanted in the substrate

4.19 Gas enhanced etching. Sample atoms interacting with sample atoms are carried away as volatile species after sputtering. In order to be effective, the beam current, dwell time, and refresh rate have to be carefully adjusted.


Focused ions
Metal atom sputtered away
Gas molecules
Dissociated species removed
Spontaneous desorption
Deposited metal atoms
Adsorbed gas molecules
Substrate
Dissociated elements remaining as impurities

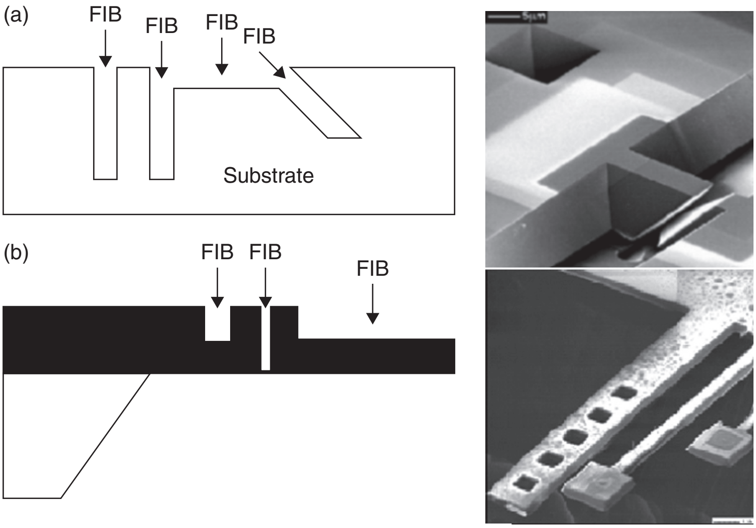
4.20 Deposition from a gas. A gas precursor delivered and adsorbed on the sample surface is dissociated by the ion beam. Volatile components of the precursor are removed, leaving a metal behind on the surface. There is a competition between the deposition process and the sputtering process.

substrate forming either volatile or non-volatile species. Evaporation removes the volatile species, while sputtering removes the non-volatile species.

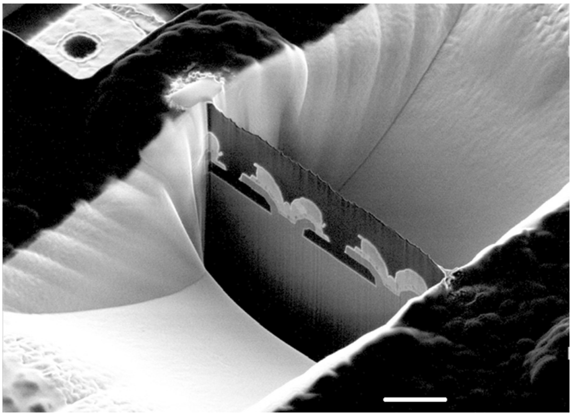
Deposition from a gas is shown schematically in Fig. 4.20. For metal deposition, typically an organo-metallic precursor compound is used, and the sequence for deposition starts with the adsorption of the precursor onto the surface, followed by ion beam induced dissociation. In this situation there is a competition between the milling of the deposited material and the deposition. The rate of deposition is a sensitive function of beam current and beam overlap frame refresh time (the time it takes to repopulate the surface after a scanned frame), both of which must be adjusted to maximize the rate (Orloff *et al.*, 2003, Chapter 6.7, Giannuzzi and Stevie, 2005; Yao, 2007, Chapter 3).

4.6 Milling

Historically, milling has been the main use of FIB. We show examples of milling for the Ga LMIS, GFIS and PFIB. As shown schematically, and by example in Fig. 4.21, FIBs can directly pattern by milling or can be used to modify existing structures.

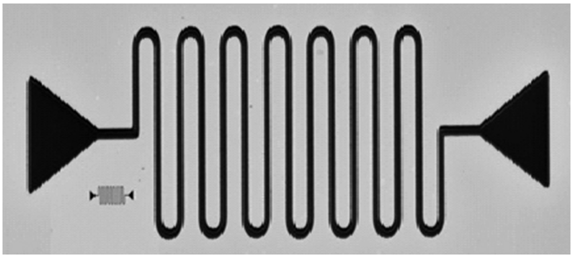


4.21 Two classes of FIB milling shown schematically and with an example of each class. (a) Milling a pattern. (b) Modifying an existing structure.

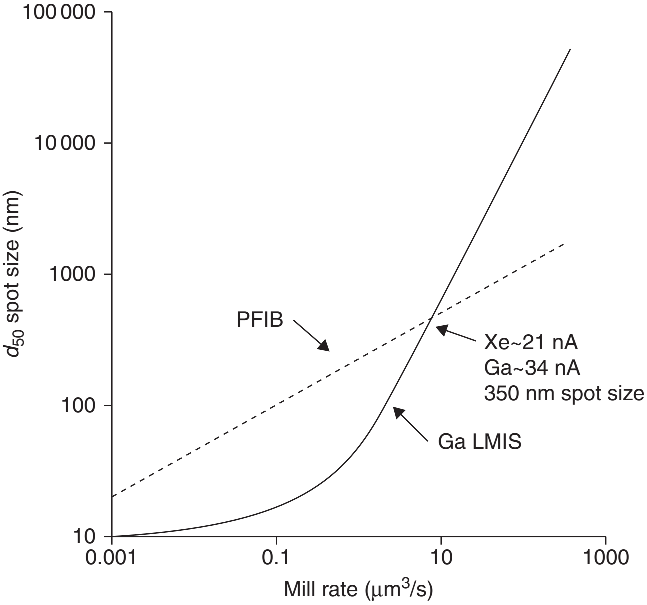


4.22 The penultimate milling step in creating a TEM sample. The next FIB step is to slice the sides and bottom of the lamella, and then to lift it out.

Presently, one of the major uses of LMIS FIB is for the preparation of TEM samples. An example is shown in Fig. 4.22. The thin lamella shown has been 'liberated' from the bulk sample by milling, and is subsequently cut at its sides and bottom by milling, and then lifted out and placed on a TEM sample grid to be viewed later in the TEM. Almost all of the process is now automated (Orloff *et al.*, 2003; Yao, 2007).

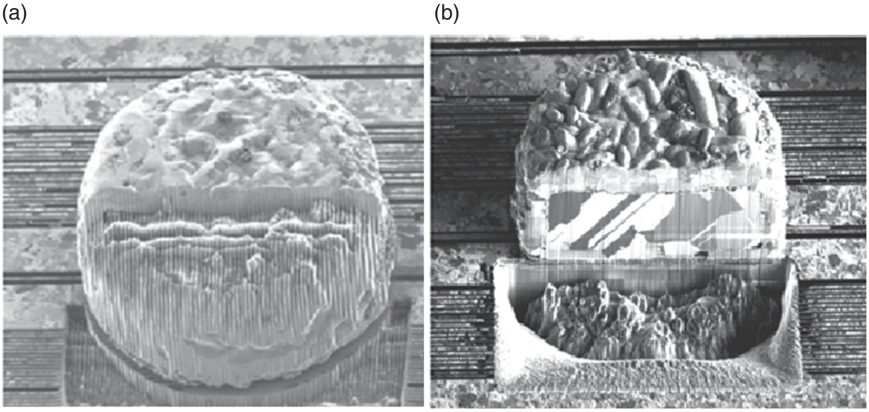


4.23 Top-down view of identically shaped micro-fluidics channels directly milled in Si at two length scales. The larger structure has a 850 μm long channel of 2 μm width milled using a 6.5 nA Ga^+ beam. The smaller structure has an 80 μm long channel of 100 nm width patterned by a 150 pA beam. Image provided by Fibics Inc., Ottawa, Ontario, Canada.



4.24 A graph comparing the d_{50} beam size diameter of Ga and Xe as a function of milling rate in Si. The milling rate scales with the beam current and nuclear stopping power. Below the crossover point near 10 μm^3/s Ga LMIS will yield a better cut. Above that point PFIB will produce smaller cuts.

Recall that the graph in Fig. 4.2 showed the size of the beam for an LMIS-based FIB system. The most useful regime for nano-machining is in the 5 pA to 10 nA interval. Above 10 nA the beam size is > 100 nm, and for large volume removal the times become onerous. Between 5 pA and 10 nA, the removal rates vary between 0.08 μm^3/min to 160 μm^3/min for common materials, so that for mill volumes ~ 10 000 μm^3 corresponding to ~22 μm × 22 μm



4.25 Example of the crossover point between LMIS and PFIB. The solder ball (a) was milled with a 20 nA 30 keV Ga^+ beam. The solder ball (b) was milled with a 30 keV, 127 nA Xe^+ beam.

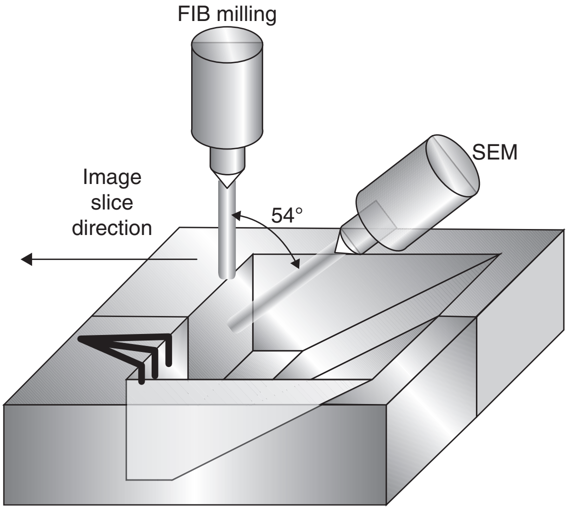
× 22 μm will take an hour using 10 nA beam current. At about 10 nA, spherical aberration starts to be the dominant aberration for LMIS-based systems, and the beam size grows quickly with increasing current. When the spot sizes for Ga- and Xe-based systems are compared as a function of milling rates, as shown in Fig. 4.24, it is seen that there is a crossover point near 350 nm beam size where the PFIB begins to dominate the LMIS-based system.

The effect of this crossover point is shown in Fig. 4.25, where on the left the ~120 μm diameter 'solder ball' from a ball grid array was milled with a 30 keV Ga^+ 20 nA beam for 30 min. The same type of solder ball on the right of the figure was completely cross-sectioned (~9 00 000 μm^3 at 500 $\mu m^3/s$) with a 30 keV 127 nA Xe^+ beam in 30 min, followed by a 23 nA 10 min. polish (to expose detail in the structure), and the image was then taken with a 80 pA Xe^+ beam.

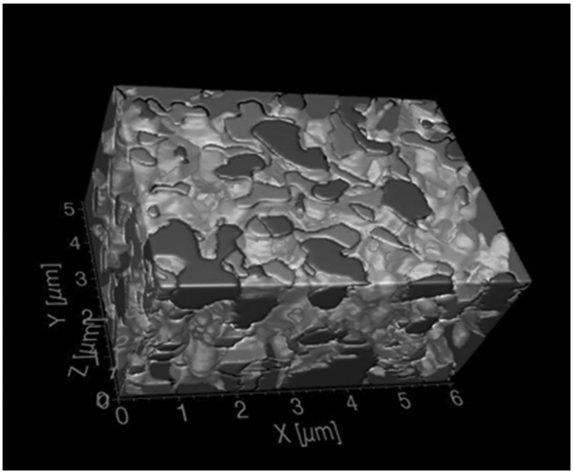
One future use of the milling capability of FIB will be in microtomography. In order to make a 3-D reconstruction of a sample, a 'slice-and view' approach is taken. Shown in Fig. 4.26 is a schematic of the process that utilizes a 'dual beam' FIB-SEM.

After the FIB has milled a new face on the sample, the SEM records and stores an image, which is later combined with previous and subsequent images to build a 3-D reconstruction of the sample, as shown in Fig. 4.27. Such reconstructions to date include both biological and inorganic samples.

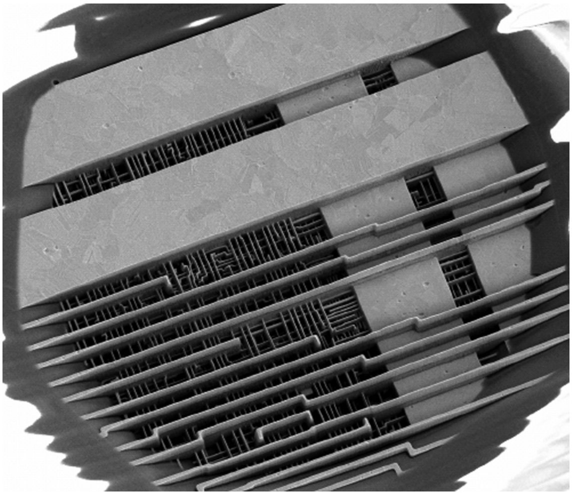
In cases where there is more than one material, gas can be used to mill selectively, as shown in Fig. 4.28 where the great etch selectivity of XeF_2 has removed Si but left intact metal structures. Such selectivity has been developed for a variety of materials (Orloff *et al.*, 2003; Chapter 6.7; Giannuzzi and Stevie, 2005; Yao, 2007, Chapter 3).



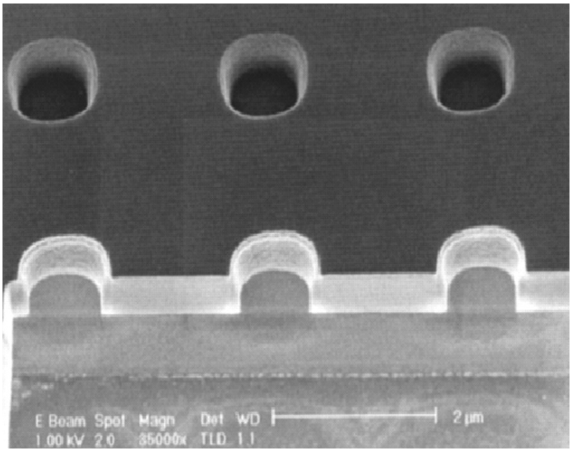
4.26 Schematic diagram of dual beam (FIB-SEM) geometry used to 'slice-and-view' a sample. The FIB mills away the exposed face of the sample which is then imaged with the SEM. The sequentially stored images are then used to reconstruct (3-D tomographically) a portion of the sample. This technique has been used in biological to petrological studies.



4.27 A three-dimensional reconstruction from the technique illustrated in Fig. 4.26. Since there are different contrast modes available in the SEM and FIB, one can imagine that the entire process could be performed in a single beam FIB (although requiring more time due to the need to tilt the sample between operations of milling and imaging).

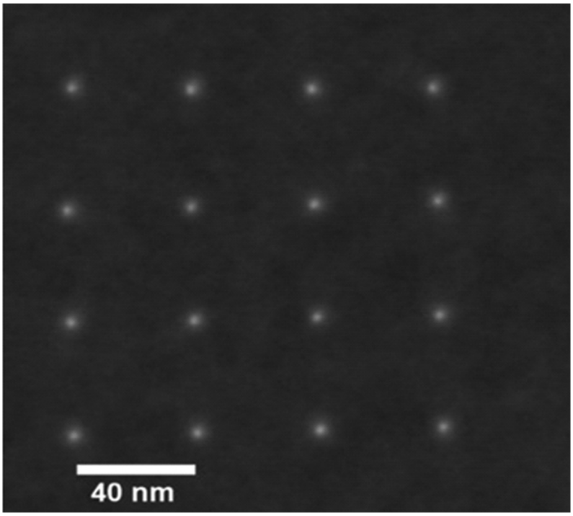


4.28 An example of selective etching. XeF_2 will selectively etch insulators while leaving most metals virtually intact. This is an example of a 'skeletal etch' process, in which a XeF_2-assisted Ga FIB was used to remove selectively the oxide of an IC device with minimal erosion of the copper metal lines.

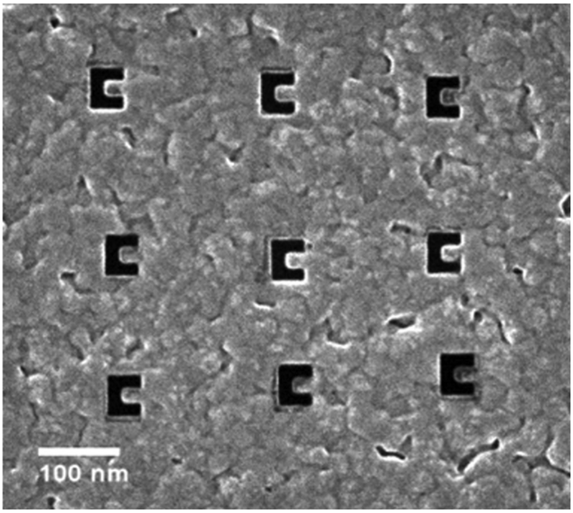


4.29 Example of water-enhanced etching to remove PMMA resist to expose the profile of the contact holes to the underlying silicon. The water can be produced as a gas in a variety of ways.

Water has been shown to be suitably efficacious in the enhanced etching of hydrocarbons while retarding the etching of metals. Figure 4.29 shows an example of water-enhanced etching to remove PMMA resist in order



4.30 An array of 4 nm diameter nanopores drilled through a 30 nm thick silicon nitride membrane. (The holes are spaced 40 nm apart.) This is a STEM image showing holes that were milled in about 1 second each. Image courtesy of Zeiss Corp.



4.31 An array of 'C'-shaped apertures milled in an aluminum film. The machining acuity was 5 nm for the corners of the apertures. A detailed examination (imaged by the same He beam that performed the milling) on the impact of aluminum grain structure from the machining is possible. This ability can aid a designer in selecting the proper material for the apertures. (Image courtesy of Zeiss Corp.)

to expose the profiles of the contact holes to the underlying silicon (Orloff *et al.*, 2003; Giannuzzi and Stevie, 2005; Yao, 2007).

The GFIS He system is capable of very fine milling due to the less destructive nature of the interaction of He as compared to Ga or Xe and the very small size of the beam. While the sputtering yields and rates are small compared to Ga and Xe, the ability to make fine structures is extended downwards on the length scale. Two examples of this fine machining are shown in Figure 4.30, where tiny holes were milled in a thin membrane with a GFIS He beam and, in Fig. 4.31, where structured shapes were machined in Al and the resultant work then imaged with the He beam.

4.7 Deposition

FIB-induced deposition of metals and insulators is a complement to milling, and also a stand-alone process for the direct production of small high resolution structures without the need of first forming a film that is etched by lithographic techniques. One of the early uses of FIB deposition was for the repair of optical and X-ray mask clear defects, where precisely defined structures were made of either carbon for optical masks and W or Au for X-ray masks. Deposition has also found use in the repair and modification of ICs, typically where fine, low-resistance metal wires and pads are deposited for connections for probing or insulators (Yao, 2007) are deposited for isolation purposes. At present, the capability exists for depositing a variety of metals and insulators. Table 4.6 shows some of the materials available for deposition.

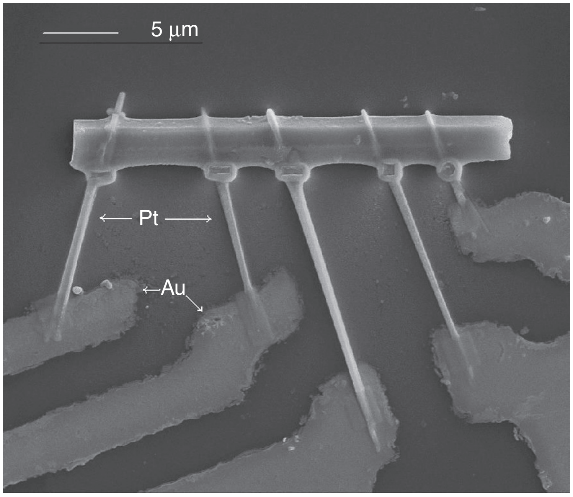
Table 4.5 Water enhancement and retardation factor for different materials

Material	Enhancement
Polyimide	18
PMMA	18
Diamond	10
Teflon	6
Au	1
SiN	0.3
SiO_2	0.3
Si	0.2
Al	

Note: Generally, the rule seems to be enhancement for hydrocarbon bearing materials, and retardation for inorganic materials.

Table 4.6 Precursor gases used for depositing metals and insulators

Deposition material	Precursor gas	Reference
Al	Trimethyl aluminum, $Al_2(CH_3)_3$	Gamo, *et al.*, 1984
Al	Trimethylamine alane	Gross, *et al.*, 1990
Al	Triethylamine alane	Gross, *et al.*, 1990
Al	Tri-isobutyl aluminum	Kubena, *et al*, 1988
W	Tungsten hexafluoride, WF_6	Gamo, *et al.* 1986
W	Tungsten hexacarbonyl, $W(CO)_6$	Stewart, *et al.*, 1989
Au	$C_7H_7O_2F_6Au$	Blauner, *et al.*, 1989
Pt	$C_9H_{16}Pt$	Tao, *et al.*, 1990
Cu	Cu(hfac)TMVS	Della Ratta, *et al.*, 1993
Ta	PMTA, $Ta(OC_2H_5)_5$	Gamo, *et al.*, 1986
Fe	Iron pentacarbonyl, $Fe(CO)_5$	Kubena, *et al*, 1988
Pa	Palladium acetate, $[Pd(O_2CCH_3)_2]_3$	Harriott, *et al.*, 1986
SiO_2	$Si(OCH_3)_4$	Komano, *et al*, 1989
SiO_2	Siloxane + O_2	Stewart, *et al*, 1995
C	Napthalene, $C_{10}H_8$	Prewitt, Mair 1991



4.32 An SEM image of an example of Ga LMIS induced deposition. A molecular organic superconductor (TMTSF) $2ClO_4$ is contacted with five Pt nano-wires deposited by FIB. (Image courtesy: Dr M.J. Naughton, Dr J.I. Oh, and Pashupati Dhakal, Boston College.)

Deposition can be limited by the rate at which gas can be delivered to the surface. As such, the technique is practical only with moderate beam currents, and is limited primarily to LMIS and GFIS (Orloff *et al.*, 2003, Giannuzzi and Stevie, 2005; Yao, 2007). A generic example of LMIS-FIB deposition is shown in Fig. 4.32.

4.8 Imaging

The impinging ion beam produces both secondary electrons and ions (Levi-Setti, 1983; Orloff *et al*, 2003). Usually, the electrons are used to form images so that the FIB acts as a microscope much like a SEM. Since the interaction of ions into samples is very different from electrons into samples (the interaction for ions that generate secondary quanta is much nearer the surface of the sample), different information is available through the FIB-generated images. Secondary-electron production is several orders of magnitude larger than the secondary-ion yield. Typically for Ga^+ ions, the secondary-ion ionization efficiency is ~ 0.001%, which is three or four orders of magnitude lower than that for secondary-electron production. Although secondary electrons are more useful for imaging, they are mostly responsible for any sample charging. Since the sample charges positively due to both the impinging positive ions and the expelled negative secondary electrons, some of the secondary electrons are attracted back to the sample, reducing the magnitude of charging but also reducing the number of secondary electrons available for imaging.

As the primary beam of ions interacts with target atom outer shell electrons, some of the electrons receive sufficient energy to be ejected from an atom. If they have sufficient energy and are within the secondary-electron escape depth λ from the surface, they can overcome the surface barrier energy and be ejected from the sample. The secondary-electron yield δ, which is the ratio of the number of electrons produced per incident ion, is given by

$$\delta = \int \frac{EBe^{-x/\lambda}}{\in R}\,\mathrm{d}x = \frac{BE}{\varepsilon R}\lambda\left(1 - e\frac{-R}{\lambda}\right) \qquad [4.8]$$

where E and R are the primary ion energy and range, ε is the energy necessary for producing a secondary electron, and B is a characteristic of the target material. λ can be estimated from

$$\lambda = 1.11\,\delta_m\left(\frac{E}{E_m}\right)^{-0.35}\left(1 - e^{-2.3\left(E/E_m\right)^{1.35}}\right) \qquad [4.9]$$

where δ_m and E_m are tabulated parameters found by experiment. Although these calculated secondary yields show promise for estimating imaging quality, they are higher than the experimental values obtained in most FIB systems; the discrepancy most likely is due to surface sample charging. Some values for insulators are given in Table 4.7.

Table 4.7 Parameters and Associated Values for Estimating Secondary-Electron Yields of some insulators

Material	B/ε	λ (nm)	R (nm)	δ_m
Si	11.5	3.2	33	5
SiO_2	16	5.8	20	12
ILD	9.3	6.2	30	8

In order to form a good image, there must be resolution, contrast, and an adequate signal-to-noise (S/N) ratio. The resolution and contrast are mostly determined by the system optics and the interaction of the ions with the sample, while the S/N is determined in a well-designed system by the length of time the beam is on the sample. With ions, a careful strategy must be employed to make a good image because of the destructive nature of the beam, so there is a fundamental limit to FIB imaging resolution, which is dictated by sputtering (Orloff *et al.*, 1996).

If an image is made with secondary electrons, in order to obtain a given S/N ratio, the ion beam must dwell at points on the sample for a time t_d so that enough secondary electrons can be collected. The number of detected electrons per pixel N_e is

$$N_e = \frac{\gamma_{eI_bt_d\eta_e}}{e} \tag{4.10}$$

where γ_e is the yield for secondary electrons, η_e is collection efficiency, and e is the charge of the electron. The average human eye requires that the minimum S/N ratio in an image be >5. In practice, there is an overlap of the beam as it is scanned along a line where negative overlap means that there is a space between adjacent positions of the beam, and positive overlap means that the beam removes material from adjacent pixels. For a beam of diameter d_b and a step size of s, the overlap OL is related to s and d_b as

$$s = d_b\left(1 - \mathrm{OL}\right) \tag{4.11}$$

The volume of material V that is removed can be written as

$$V = \frac{YAN_{\mathrm{Ga}}}{N_A\left(1 - \mathrm{OL}\right)^2} \tag{4.12}$$

where N_{Ga} is the number of incident ions per pixel and N_A the atomic mass. It can be shown (Orloff *et al.*, 2003) that the S/N and N_{Ga} are related by

Table 4.8 Sputtering can limit the resolution available in an image. Shown are minimum resolution limits d for a variety of materials for images with *S*/*N* = 20 and *S*/*N* = 5. Even though the resolution is higher for *S*/*N* = 5, it may be more difficult to distinguish an object due to the noise level. *Y* is the sputter yield

	C	Al	Si	Ti	Cr	Fe	Ni	Cu	Mo	Ta	W	Au
Y	2.7	2.9	2.2	3.3	1.2	3.9	2.0	3.4	1.3	2.8	1.2	14
$d_{S/N=20}$	4.7	5.5	5.3	5.9	3.8	5.5	4.3	5.2	4.1	5.7	4.1	9.5
$d_{S/N=5}$	1.9	2.2	2.1	2.4	1.5	2.2	1.7	2.1	1.6	2.3	1.6	3.8

$$N_{\mathrm{Ga}}=\left(\frac{S}{N}\right)^2\frac{(1+\delta)}{\delta} \tag{4.13}$$

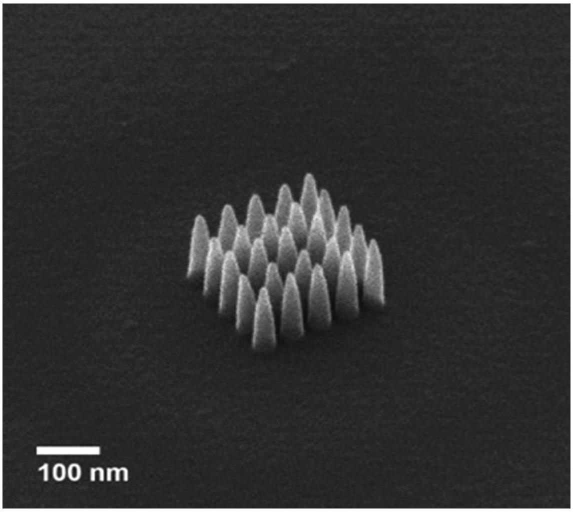
then

$$V=\frac{YA(S/N)^2(1+\delta)}{\rho N_A(1-\mathrm{OL})^2\delta} \tag{4.14}$$

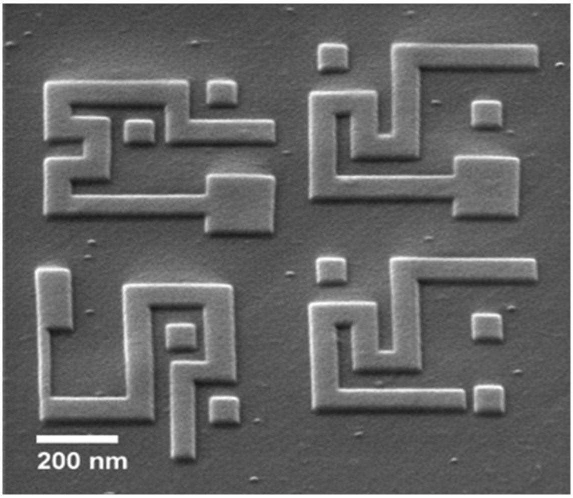
Assuming that the characteristic size d of a particle of volume *V* is $V^{1/3}$, then the minimum resolution can be calculated, assuming it is d. It is clear from (Equation [4.14]) that the limiting parameter for image formation not under the control of the user is the sputter yield *Y* for a given incident ion and sample. From experience on a variety of samples, in order to achieve an *S*/*N* ratio of 20 (a very nice image), it has been found that the number of incident ions must be ~ 700 ± 200. For most eyes, the minimum acceptable *S*/*N* is five. There is a competition between the rate at which material is removed via sputtering and the collection of secondary electrons that can be collected. For high sputter rate materials, not as much information can be gathered. Table 4.8 gives some values of detectible resolution for a variety of materials at different *S*/*N*.

It is clear that if less *S*/*N* is demanded, then the resolution gets better. However, it becomes harder to discern what is seen. Also, it must be noted that, for extended structures that have sufficient contrast, the above analysis does not hold, and that the resolution that can be achieved can be higher. In the He GFIS system, the virtual lack of sputtering and the small virtual source size allows for very high imaging resolution. Examples of FIB images are shown below. In Fig. 4.35 it is clear that FIB gives different information than SEM.

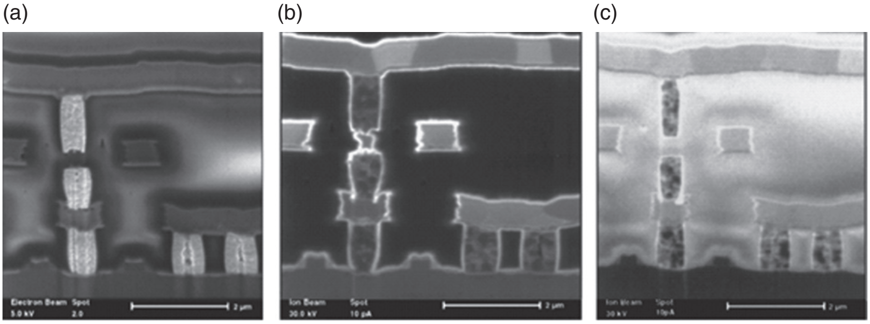
That very high resolution can be achieved is shown in Fig. 4.36, where an image from a He GFIS system of an uncoated biological sample is shown.



4.33 An array of nano-pillars created by GFIS He beam induced chemistry. These pillars were fabricated using a platinum-bearing precursor gas. The size (35 nm diameter) and tight array pitch (50 nm) shows the deposition capabilities of He beam induced processing. (Image courtesy of Zeiss Corp.)



4.34 A test pattern of conducting lines and pads produced by GFIS He ion beam induced deposition from platinum precursor gas. Tight spatial control is evident as seen in the sharp corners and delineation of the deposited features. (Image courtesy of Zeiss Corp.)

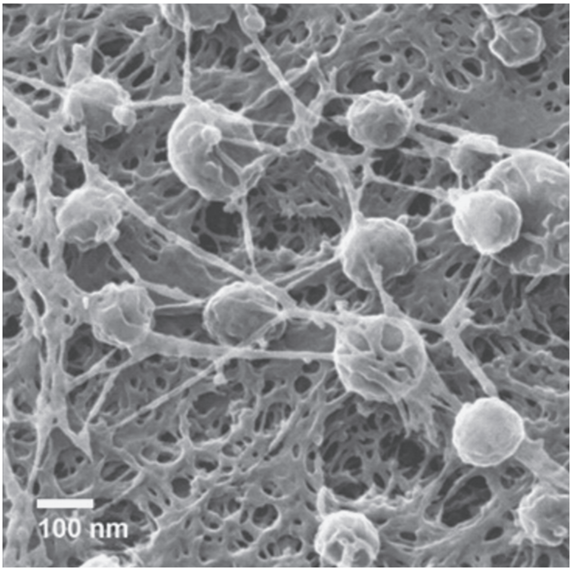


SEM image (secondary electrons)

Scanning ion image (secondary electrons)

Scanning ion image (secondary ions)

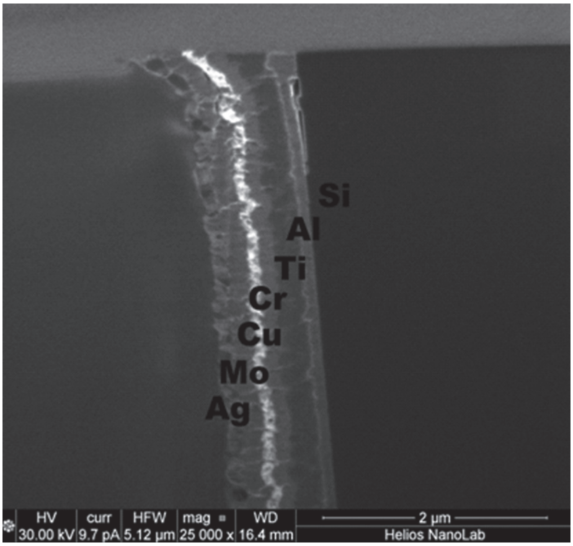
4.35 The differences in contrast mechanisms available. (a) SEM image with secondary electrons, (b) scanning ion image made by collecting secondary electrons, and (c) scanning ion image made by collecting secondary ions. (Courtesy of P. Carleson, FEI Co.)



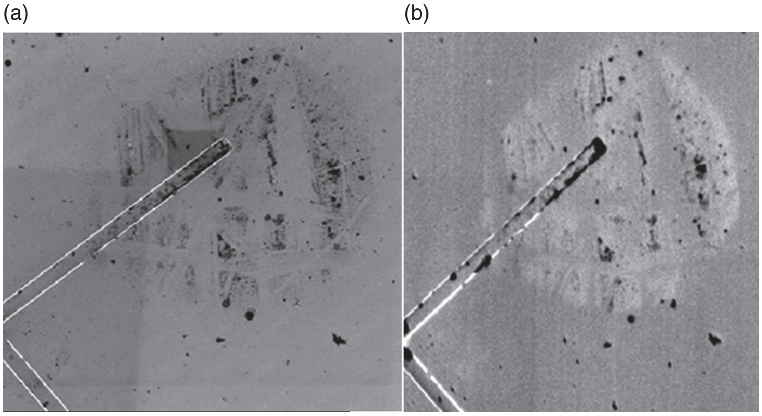
4.36 Cell body exposed to a virus. The membrane has been stripped away in the sample preparation process, showing the degree to which the virus particles have infiltrated the cell. Uncoated sample. (Image courtesy of Zeiss Corp.)

Different materials yield different contrast levels, as can be seen in Fig. 4.37.

Resolution is not the only characteristic of an image that yields information. Contrast also allows the discernment of objects, and can arise by topography changes or material differences.


Si
Al
Ti
Cr
Cu
Mo
Ag
HV
30.00 kV
curr
9.7 pA
HFW
5.12 µm
mag
25 000 x
WD
16.4 mm
2 µm
Helios NanoLab

4.37 A layered sample designed to show the relative elemental contrast of several common materials. The image was made by collecting secondary electrons. It is interesting to note that the brightness of the layers in the image correlates also with their sputter yields. (Image courtesy of Dr Lucille Giannuzzi, Giannuzzi & Associates.)



4.38 Two images of the same area showing the contrast differences between electrons and Li ions. The ion image was made in a MOT-based CAIS system. (Courtesy Chad Rue, Fei Co.)

Figure 4.38 shows a comparison of images of the same region of a sample between a Li MOTIS ion source and an SEM, where it can be seen that the contrast modes are different.

4.9 Spectroscopy

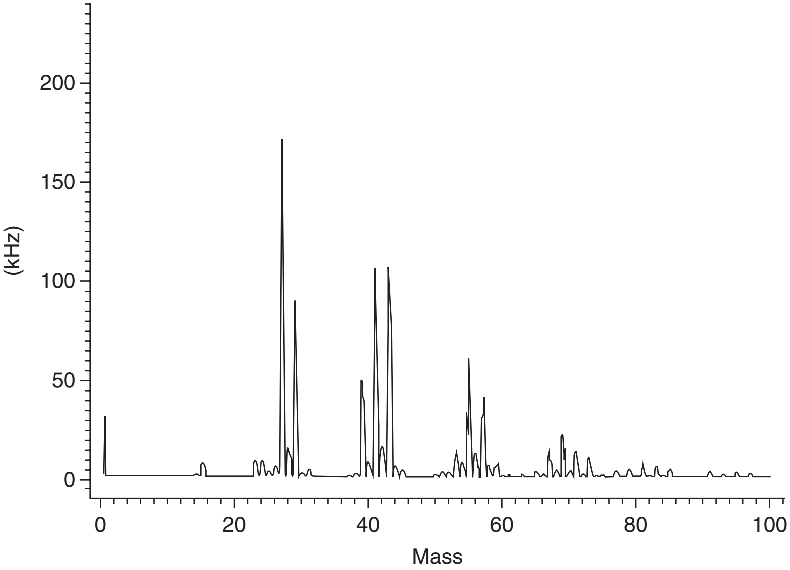
Since some fraction (~0.1%) of the sputtered atoms from the sample is ionized, they can be collected and mass analyzed, so that the destructive properties of the FIB can be advantageously used for the chemical analysis of surfaces. It is possible to control the rate of milling caused by the primary beam, so that spectra or either two- or three-dimensional chemical maps of the sample can be made. Due to the high spatial resolution inherent in FIBs, the chemical maps obtained are also of high lateral spatial resolution and, under suitable conditions, the depth resolution can also be very high. Usually in FIB systems, collecting secondary ions to determine target composition at nanometer scales is accomplished with SIMS. There are three main types of FIB/SIMS systems: (1) RF quadrupole, (2) time-of-flight (TOF), and (3) magnetic sector. The RF quadrupole and magnetic sector use a standard FIB column to form a probe, which is scanned across the sample collecting secondary ions and transporting them to either a RF quadrupole or magnetic sector, where the ions are filtered and detected in a serial mode as the mass channel is swept through its range while the primary beam continuously strikes the sample. The TOF technique requires that a mono-isotopic source material to be used, and that the primary beam be pulsed (turned off–on–off-…) for about 5 ns to allow the ions time to travel the length of the detector.

Much of the pioneering work coupling a Ga LMIS-based source with SIMS was performed at The University of Chicago by Levi-Setti (Chabala *et al.*, 1992, 1994, Levi-Setti *et al.*, 1994). In relation to other surface analytical techniques, several trade-offs must be considered. While the lateral and depth resolution are high (20 nm laterally, and 1 nm depth) with a minimum detectable concentration as good as a few hundred parts per million at a signal-to-noise $>10^6$ with isotope sensitivity, usually quantification is difficult due to large variations in sputtering and ionization yields that are sensitive to surface oxides, which greatly enhance the ion yield; and since the sample is consumed as it is being analyzed, no reexamination is possible, so that parallel collection techniques must be used in certain cases. In addition, primary ions are implanted into the sample during the analysis and may react with it or confuse the interpretation of data.

Spectra can be obtained, as shown in Fig. 4.39, and also high resolution elemental maps of the spatial distribution of sample components can be made as shown in Fig. 4.40

4.10 Conclusion and future trends

'Prediction is very hard, especially about the future' (Berra, 1997). There is little doubt that FIB will continue to be used for failure analysis, circuit edit, and for device and IC restructuring and modification in the semiconductor



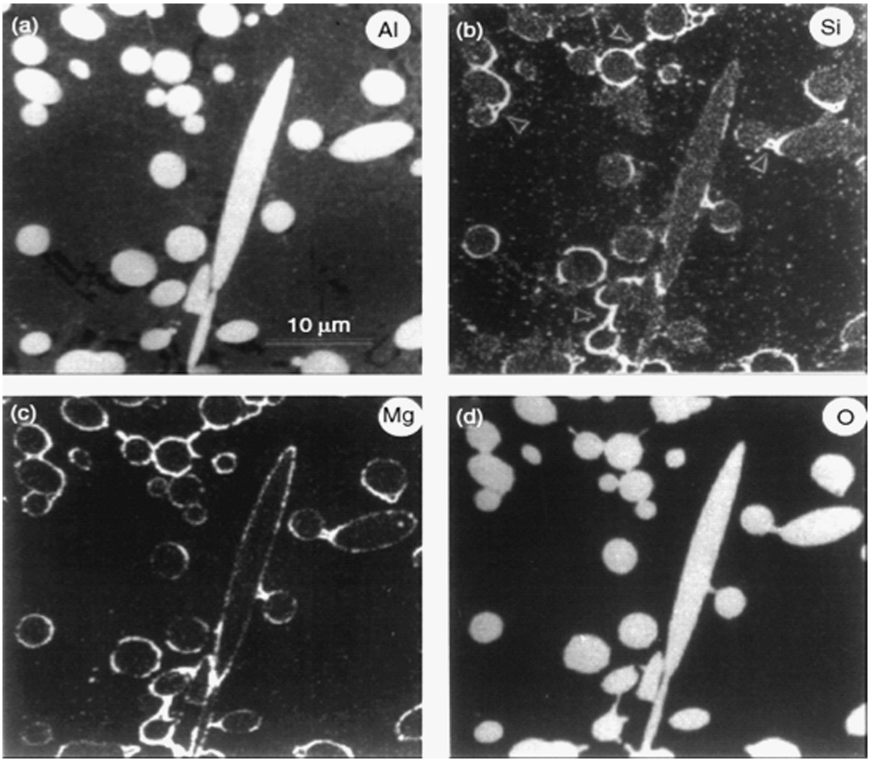
4.39 Mass spectrum taken with a Ga^+ beam. This spectrum was obtained by collecting the secondary ions and passing them through a quadrupole mass analyzer.

industry, fueled by the Information Age need for transistors (the most numerous thing humans make). LMIS-based FIB now plays an essential role there, and as node dimensions continue to decrease, both the GFIS and CAIS will be utilized. The role for PFIB for large volume milling should also increase in the future, especially in cases where time is critical.

Many new technologies take the better part of 30 years to home-in on a useful niche. Certainly GFIS sources, and probably the CAIS sources, will find future utilization as design sizes shrink, and both machining and imaging requirements will increase the need for higher resolution solutions. It may be that FIB will finally be utilized in biology.

There seems to be a trend for marrying technologies, such as FIB, SEM, and lasers, where in one venue different tasks can be accomplished without the need to transport a sample between machines. Such systems that also employ gas-chemistry solutions should continue on into the future as more techniques evolve to perform tasks of selective and accelerated milling and deposition. The unique and complementary imaging capabilities of FIB should make FIB valuable for future work.

As femtosecond laser ablation becomes more known and used (with a million times faster rate of material removal as compared to FIB), it should offer some competition to PFIB, but might still require some sort of FIB polishing capability.



4.40 SIMS elemental maps of the end region of the X chromosome of the fruit fly *D. melanogaster*. The maps are produced by scanning the beam over the sample and collecting secondary ions and passing them through a spectrometer. There are simple techniques to make maps of regions using different masses to provide image contrast. Left image CN^- map. Right image Br^- map. (Courtesy of R. Levi-Setti,, The University of Chicago.)

FIB systems will continue to become more automated and reliable. They will find their way into most large industries (across length scales ranging from mining, oil and gas to MEMS, the nano-world, and biology) typically as multifunctional dual beam systems. One can imagine eventually that MEMS and nanotechnology will be researched and developed in '*laboratories-in-a-can*,' of which FIB will be a participant and that FIB will be used as a 'nano-scalpel' in biology.

4.11 References

Benninghovan, A., Rudenauer, F. G. and Werner, H. W. (1987). *Secondary Ion Mass Spectrometry*, Wiley, New York.

Berra, Yogi, (1997). *The Yogi Book: I Really Didn't Say Everything I Said!*, Workman Publishing, New York.

Chabala, J. M., Levi-Setti, R., Li, L., Parker, N. W. and Utlaut, M. (1992). Development of a magnetic sector-based high lateral resolution scanning ion probe, In: A. Benninghoven, K. T. F. Janssen, J. Tumpner, and H. W. Werner, (Eds.), *Secondary Ion Mass Spectrometry, SIMS VIII*, Wiley, Chichester, 179.

Chabala, J. M., Soni, K. K., Li, J., Gavrilov, K. L. and Levi-Setti, R. (1994). High-resolution chemical imaging with scanning ion probe, *SIMS. Int. J. Mass Spectr. Ion Proc.*, **143**, 191–212.

Chandrasekhar, S. (1943). Stochastic problems in physics and astronomy, *Rev. Mod. Phys.*, **15**, 1–89.

Crewe, A. V. (1973). Production of electron probes using a field emission source, In: *Progress in Optics*, Vol. XI, E. Wolf, Ed., North-Holland, Amsterdam, Chapter 5.

Crewe, A. V. (1987). Optimization of small electron probes, *Ultramicroscopy*, **23**, 159–168.

Dahl, D. A., Delware, J. E. and Appelhans, A. D. (1990). SIMION PC/PS2 electrostatic lens program, *Rev. Sci. Instrum.*, **61**, 607–609.

Giannuzzi, L. A. and Stevie, F. A., (Eds.), (2005). *Introduction to Focused Ion Beams:Instrumentation, Theory, Techniques and Practice*, Springer, Berlin.

Grivet, P. (1972). *Electron Optics*, 2nd ed., Pergamon Press, Oxford.

Hill, A. R. (1968). Uses of fine focused ion beams with high current density, *Nature*, **218**, 262.

Jiang, X. R. and Kruit, P. (1995). Influence of lens aberrations and Coulomb interactions on choice of magnification and aperture size in a two lens focused ion beam system, *J. Vac. Sci. Technol.*, **B14**, 1635–1641.

Kellogg, S., Schampers, R., Zhang, S., Graupera, A., Miller, T., Laur, W. D. and Dirriwachter, A. (2010). High throughput sample preparation and analysis using an inductively coupled plasma (ICP) focused ion beam source, *Microsc. Microanal.*, **16**, 222–223.

Kingham, D. R. and Swanson, L. W. (1984). Shape of a liquid metal ion source, *Appl. Phys. A*, **34**, 123–132.

Krohn, V. E. and Ringo, G. R. (1975). Ion source of high-brightness using liquid metal, *Appl. Phys. Lett.*, **27**, 479–481.

Levi-Setti, R. L. (1974). Proton Scanning microscopy: feasibility and promise, *Scanning Electron Microsc.*, **1**, 125.

Levi-Setti, R. (1983). Secondary electron and ion imaging in scanning ion microscopy, *Scanning Electron Microsc.*, **1**, 1–12.

Levi-Setti, R., Chabala, J., and Smolik, S. (1994). Nucleotide and protein distribution in BradU-labelled polytene chromosomes revealed by ion probe mass spectrometry, *J. Microsc.*, **175**(1), 44–53.

Melngailis, J. (1987). Focused ion beam technology, *J. Vac. Sci. Technol.*, **B5**, 469.

Melngailis, J. (1991). Focused ion beam lithography and implantation, In: W. Glendinning and J. Helbert, (Eds.), *Handbook of VLSI Microlithography*, Noyes Publications, Park Ridge, NJ.

Munro, E. (1973). *Image Processing and Computer-Aided Design in Electron Optics*, P. W. Hawkes, (Ed.), Academic Press, London.

Notte J. and Ward B. (2006). Sample interaction and contrast mechanisms of the helium ion microscope. *Scanning*, **28**, 13–16.

Orloff, J. (1993). High-resolution focused ion beams, *Rev. Sci. Instrum.*, **64**(5), 1105–1130.

Orloff, J. and Swanson, L. W. (1975). Study of a field ionization source for microprobe applications, *J. Vac. Sci. Technol.*, **12**, 1209–1213.

Orloff, J., Utlaut, M. and Swanson, L.W. (1996). Fundamental limits to imaging resolution of focused ion beam systems, *J. Vac. Sci. Tech*, **B14**, 3759–3664.

Orloff, J., Utlaut, M. and Swanson, L. W. (2003). *High Resolution Focused Ion Beams*, Plenum, New York.

Reijnders, M. P., van Kruisbergen, P. A., Taban, G., van der Geer, S. B., Mutsaers, P.H.A., Vredenbregt, E. J. D. and Luiten, O. J. (2009). Low-energy-spread ion bunches from a trapped atomic gas, *Phys. Rev. Lett.*, **102**, 034802–034805.

Seliger, R. L. (1972). E × B Mass-separator design, *J. Appl. Phys.*, **43**, 2352–2357.

Seliger, R. L., Ward, J. W., Wang, V. and Kubena, R. (1979). A high-intensity scanning ion probe with submicrometer spot size, *Appl. Phys. Lett.*, **34**, 310.

Sigmund, P. (1969). Theory of sputtering I: sputtering yield of amorphous and polycrystalline targets, *Phys. Rev.*, **184**, 383–416.

Sigmund, P. (1977). Sputtering processes: collision cascades and spikes, In: N. H. Tolk, J. C. Tully, W. Heiland, and C. W. White, (Eds.), *Inelastic Ion-Surface Collisions*, Academic Press, New York.

Smith, N. S., Skoczylas, W. P., Kellogg, S. M., Kinion, D. E., Tesch, P. P., Sutherland, O., Aanesland, A. and Boswell, R. W. (2006). High brightness inductively coupled plasma source for high current focused ion beam application, *J. Vac. Sci. Technol. B*, **24**(6), 2902–2906.**A**

Stewart, D., Stern, L., and Morgan, J. (1989). Electron beam x-ray and ion beam technologies: submicrometer lithographies VIII, *Proc. SPIE*, **1089**, 18–25.

Swanson, L. W. (1994). Use of the liquid metal ion source for focused ion beam applications, *Appl. Surf. Sci.*, **76/77**, 80–88.

Tandare, V. N. (2005). Quest for high brightness, monochromatic noble gas ion sources, *J. Vac. Sci. Technol. A*, **23**, 1498–1507.

Ward, J. W., Kubena, R. L., and Joyce, R. J. (1991). An ion counting apparatus for studying the statistics of ion emission from liquid metal ion sources, *J. Vac. Sci. Technol. B*, **9**(6), 3090–3094.

Ward, J. W., Kubena, R. L. and Utlaut, M. (1988). Transverse thermal velocity broadening of focused beams from liquid metal ion sources, *J. Vac. Sci. Technol. B*, **6**(6), 2090–2094.

Ward, J. W., Utlaut, M. and Kubena, R. L. (1987). Computer simulation of current density profiles in focused ion beams, *J. Vac. Sci. Technol. B*, **5**(1), 169–174.

Yao, N. (2007). *Focused Ion Beam Systems*, Cambridge University Press, Cambridge, U.K.

Ziegler, J. F., Biersack, J. P. and Littmark, U. (1991). *Computer Code TRIM-91*. IBM-Research, Yorktown, New York.

5

Masks for micro- and nanolithography

E. GALLAGHER and
M. HIBBS, IBM Microelectronics Inc., USA

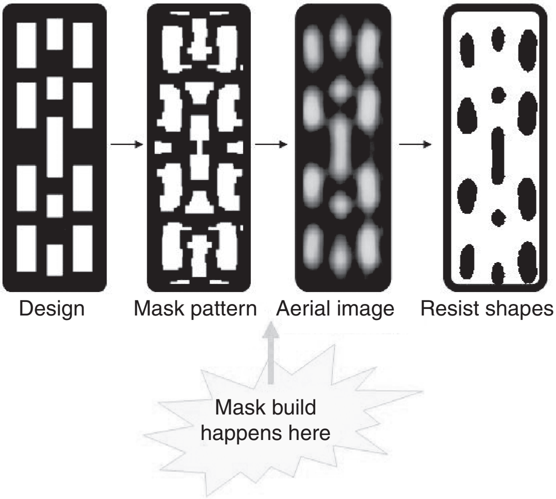
DOI: 10.1533/9780857098757.158

Abstract: This chapter provides a general introduction to the masks used in microlithography. The mask contains the master copy of the pattern that is printed on wafer to create a semiconductor chip. Because this master copy is integral to the fabrication of every semiconductor chip of a given design, the mask must be tightly controlled for the ultimate semiconductor device to function properly. Mask types, mask materials, and their applications in semiconductor manufacturing are reviewed. The specific case of optical photomasks is used as an example to enable a clear overview of mask materials, processing, metrology, and defect control strategies.

Key words: mask, photomask.

5.1 Introduction

Modern semiconductor chips contain billions of circuit elements connected by miles of conductive connections. These structures are laid down in many dozens of discrete layers, each layer defining the design of a wiring pattern, interconnection layout, or transistor array. The huge number of pattern elements in each layer requires some sort of printing technology to mass-produce the patterns across the wafers that are later diced into chips. This technology is known as microlithography. An integral part of microlithography is the master template, known as the mask. Figure 5.1 illustrates where mask fabrication fits into the flow that transforms the designer's intent into a resist image on the wafer, and how the patterns are altered during that process. Once the image is formed on the wafer, the pattern is etched into the surface of hundreds of chips arrayed across the wafer's surface and becomes part of the semiconductor device. The mask contains the master copy of the pattern being printed, and is a simple, passive, device in concept. This chapter describes the not-so-simple reality of building a lithographic mask, including an overview of the processes and exacting requirements.



5.1 Mask data preparation and processing are intermediate steps in the transformation of the designer's intent into a wafer resist image. The first image is the design data. The second image is the mask data that have been modified to accommodate for distortions from exposure, resist, and mask effects. The mask is patterned and built using the modified data. The third image is an aerial exposure of the completed mask. The final threshold model of the aerial image shows an approximation of the wafer shapes that will result.

5.1.1 Mask types

The pattern contained on the mask must be transferred to the wafer. The most common method of transfer is to shine light through the mask and create an image in a photosensitive layer on the wafer. Some regions of the mask transmit, while the others absorb the light. The generic term "mask" is narrowed to the more specific "photomask" when it is imaged using exposure in the wavelength range of 13.5–436 nm. "Reticle" is sometimes used interchangeably with "photomask," but it is better avoided because of its complex history and ambiguous meaning (Sze, 1983). Table 5.1 summarizes the most common types of exposure masks with defining features that are described below.

Photomasks require a stable substrate material to act as a carrier for the pattern-containing absorber film. Glass is the most common substrate material. It has many desirable qualities; it is abundant, has high transmission at exposure wavelengths, and can be polished to very tight tolerances.

Chromium has been the absorbing material of choice for the opaque film for several decades because of its good adhesion to glass, chemical stability, and mechanical durability. The phrase "chrome on glass" (COG) has

Table 5.1 Summary table of mask types

Mask type	Exposure source	Substrate	Absorber	Transmissive or reflective
Optical binary	Visible or UV light	Glass or fused silica	Chromium, molybdenum silicide	Transmissive
Optical phase shifting	Visible or UV light	Glass or fused silica	Molybdenum silicide	Transmissive
X-ray proximity	1 keV X-ray	Polymer or silicon membrane	Gold, tungsten	Transmissive
Electron projection	10–100 keV electrons	Polymer or silicon membrane, stencil	Gold, tungsten	Transmissive
Ion projection	50–250 keV ions	Silicon stencil	Gold, tungsten	Transmissive
EUV	13.5 nm soft X-rays	Low expansion glass or ceramic	Tantalum nitride	Reflective

been coined for this type of mask. Typically, COG masks block the light where absorber is present and transmit light where the absorber has been removed. Because of this two-state nature, COG masks are also called binary masks. Another type of photomask uses the same transparent substrate, but a deliberately leaky absorbing film. If the absorber is allowed to transmit a small fraction (typically 6%) of the incident light, and the optical phase of this transmitted light is shifted by 180°, then the image contrast will be enhanced. This type of mask is called an attenuating phase shift mask.

Significant effort has been invested in developing both X-ray and ion beam lithography, but neither technology has been used in manufacturing. Both X-rays and ion beams are heavily absorbed by almost all materials, but can be transmitted through thin polymeric membranes. A layer of high-atomic number material, such as gold or tungsten, is used as the patterning layer on the thin membrane to create a mask. Even the thin membranes are not transmissive enough for very low energy electrons or heavy ions; for these applications stencil masks must be used. Stencil masks are similar to membrane masks except that the transparent regions are etched all the way through the membrane, leaving a clear path for the charged particles to pass through. To maintain the mechanical integrity of the membrane, each opening in the mask must be very small. Large design features must be created by stitching together patterns from multiple masks containing complementary portions of the design.

5.1.2 Mask use

The masks can be used in different modes. For contact or proximity use, the mask is held either in contact with the wafer surface or in close proximity to it. Both of these modes of operation have severe drawbacks. Contact printing inevitably contaminates the mask with photoresist or debris from the wafer surface, and proximity printing suffers from poor resolution unless the proximity gap is impractically small or the exposure wavelength is very short. A far more common type of exposure is projection mode, in which the mask image is projected onto the wafer surface through a lens. This lens can be a conventional lens for visible or ultraviolet light, albeit with extraordinary requirements for low distortion, low field curvature, and low aberrations. It is also possible to use electromagnetic lenses for focusing charged particle beams onto the chip. Projection lithography protects the mask from damage and has another major advantage: the image on the chip may be smaller than the corresponding mask image. This image size reduction allows the mask features to be made larger than the final dimension of the wafer features, and greatly eases the tolerances required for the mask features. The most common reduction factors in optical projection lithography are 5× and 4×, but reduction factors as large as 10× and as small as 1× have been used.

Very recently masks have been developed for extreme ultraviolet (EUV) wavelengths around 13.5 nm. This wavelength is much shorter than the wavelengths used for optical photomasks, but much longer than X-ray wavelengths. Masks and exposure lenses designed for EUV use a unique reflective design. A multilayer stack of molybdenum and silicon films has been developed with a high reflectivity at 13.5 nm wavelength. This multilayer mirror is coated with a thin absorbing film, which is patterned just as a conventional mask absorber would be. Because the substrate does not have to transmit the EUV wavelength, it is selected solely for its mechanical and thermal stability.

Optical and ultraviolet (UV) masks are used with a variety of exposure wavelengths. Because the refractive lenses used in semiconductor microlithography are weakly corrected for chromatic aberration, the light sources are usually monochromatic, such as atomic spectral lines from arc lamps or lasers. Table 5.2 gives a list of wavelengths used for semiconductor lithography, and their sources. For fundamental physical reasons, the ultimate resolution of a lens is proportional to the wavelength of the light source. Consequently, there has been a steady reduction of wavelengths from the earliest days of semiconductor lithography in the 1970s to the present. During this time period, the minimum dimension of circuit patterns created on the wafer has shrunk from a few microns to less than 25 nm.

Table 5.2 Lithographic wavelengths and light sources

Wavelength (nm)	Spectrographic line	Light source	Comments
436	g-line	Hg-Xe arc lamp	Introduced late 1970s
405	h-line	Hg-Xe arc lamp	Used infrequently since early 1980s
365	i-line	Hg-Xe arc lamp	Introduced mid 1980s
248	—	Kr-F excimer laser	Introduced late 1980s
193	—	Ar-F excimer laser	Introduced early 1990s
157	—	F_2 laser	Oxygen-free atmosphere required
13.5	—	Sn or Xe plasma	Vacuum environment required

The following sections will describe many aspects of masks in more detail. There are sections on materials, the fabrication process, pattern metrology, and defect inspection and mitigation. Since photomasks dominate all aspects of the industry today, these topics will be discussed largely in the context of photomasks.

5.2 Mask materials

There are two principal components to a mask: a strong, rigid substrate which provides a stable platform, and a thin absorber film which provides a high degree of contrast for the etched mask pattern. There are several materials that can be used for these two components, and these are discussed in the following sections.

5.2.1 Mask substrate materials

The substrate used for optical masks has several stringent requirements. It must be rigid, transparent to the exposure wavelength, and thermally stable. It also must be resistant to damage such as darkening from the light used for the exposure. Table 5.3 summarizes different substrate materials and their most relevant properties (Davila, 2010). Masks used for exposure wavelengths longer than 400 nm often use borosilicate crown glass (BSC), which is transparent down to 380 nm wavelength and has a moderately low coefficient of thermal expansion. Masks used for shorter wavelengths have universally been made of fused silica, which is transparent to 180 nm. Because fused silica has good optical properties for all wavelengths of lithographic interest, and because its stiffness and coefficient of thermal

Table 5.3 Mask substrate materials

Material	Coefficient of thermal expansion (ppm/K)	Young's modulus (GPa)	Elastic limit (MPa)	Optical transmission range (nm)
Soda-lime glass	8.9	72	40	330–2700
Borosilicate crown glass (BSC)	3.3	65	50	380–2500
Fused silica (Corning, 1991)	0.52	73	50	180–2500
Fluorinated fused silica (Hosono, 1991)	–	–	–	155–2500
Calcium fluoride (Batchelder, 1964; Dickinson, 1975; Anon, 1982)	18.9	76	36	140–7000
LTEM (Corning, 2006)	0.00 ± 0.03	68	50	N/A

expansion are superior to BSC, it has largely replaced BSC for high-quality masks used for semiconductor lithography. Some non-demanding lithographic applications can use masks made of soda-lime glass, which is much cheaper than BSC or fused silica, but has a high coefficient of thermal expansion.

For a period in the late 1990s and early 2000s there was a burst of interest in the 157 nm exposure wavelength, due to the availability of a fluorine laser light source at that wavelength. Fused silica was far too opaque to be used as a mask substrate for 157 nm, and other materials such as calcium fluoride were considered for a while. It was quickly found that an extremely high coefficient of thermal expansion made this material unsuitable. During the same period, it was found that the lower wavelength limit of fused silica transmission could be reduced below 157 nm by heavily doping the material with fluorine, enabling its use for 157 nm masks. This technology was dropped from device makers' lithography roadmaps by the mid 2000s because the marginal resolution benefit over existing 193 nm options was not sufficient to offset the large investment required. The shortest wavelength in common use for optical or UV lithography is now 193 nm, which does not require the use of fluorinated fused silica.

EUV mask substrates do not have to be transparent, since EUV masks are used only in reflection. This allows the use of a low thermal expansion material (LTEM), which has a coefficient of thermal expansion more than an order of magnitude lower than fused silica. This type of material is typically made of fused silica doped with a percentage of titanium dioxide, and contains a mixture of crystalline and amorphous phases. Various suppliers have marketed the material using their own names: Zerodur from Schott AG and ULE from Corning.

There are no fundamental requirements that dictate the size of a photomask, but practical considerations of lens field sizes and strength and rigidity requirements for the mask have led to adoption of a few standardized sizes. During the mid 1990s the industry standardized mask sizes and magnification. The most common size for masks used in semiconductor lithography became 6″ × 6″ × 0.25″ (152.4 mm × 152.4 mm × 6.35 mm). Another standard size in use for less critical applications is 5″ × 5″ × 0.090″, but the thinner masks do not have enough rigidity for the most demanding lithographic applications. Masks for specialized applications (such as flat-panel displays) can be built with greater dimensions than one meter square.

5.2.2 Requirements for fused silica mask substrates

Fused silica (amorphous silicon dioxide) is a material frequently used for optical applications from the ultraviolet to the infrared. Fused silica is nominally made of pure silicon dioxide, but unavoidable impurities can have an effect on the optical performance of a photomask. Low-grade fused silica, sometimes known as fused quartz, is made from natural sand deposits and has high levels of metallic impurities. These impurity concentrations may be as high as 10–20 ppm and cause a strong blue to green fluorescence when exposed to UV light below 290 nm wavelength. This fluorescence is undesirable because it introduces stray light outside the desired band of exposure wavelengths and reduces the optical transmission of the mask. Synthetic fused silica is made from chemically purified feedstock with less than 1 ppm of metallic impurities and negligible levels of fluorescence. Besides metals, the principal impurity in fused silica is the hydroxyl group (OH). Depending on manufacturing methods, OH content can vary from a few ppm to over 1000 ppm. This impurity strongly affects the transmission at infrared wavelengths, but has very little effect on UV transmission. In fact, it seems that OH content protects fused silica from short-wavelength radiation damage. Fused silica is susceptible to the formation of light-absorbing color centers after long exposure to UV light, and OH groups suppress the formation of E’ color centers (Yamagata, 1992).

For fused silica of sufficient purity, the bulk absorption at 193 or 248 nm is nearly negligible. There are transmission losses due to reflection at the two surfaces of the fused silica, but they are small, in the range of 3.5–5% per surface. These reflective losses could be reduced by using anti-reflective coatings on one or both mask surfaces, but this is not usually done.

The pattern-bearing surface has extremely stringent requirements. No particulate contaminants, pits, or other surface imperfections larger than a few nanometers are permitted. The surface must also be extremely flat, so that all parts of the pattern are near the center of the focal range of the projection optics. It is relatively easy for mask blanks to be made with the critical surface flatter than a few hundred nanometers, but the initial flatness of the fused silica substrate can be degraded if the deposited absorber films have high levels of stress.

Internal stress within the fused silica substrate does not significantly affect optical transmission or mechanical stability, but it does induce an optical phenomenon called stress birefringence. This has the effect of changing the refractive index of the fused silica as a function of the direction of the incoming light's polarization. If the mask is used with unpolarized illumination, then birefringence in the mask substrate has no effect. Polarized light, however, is beginning to be used with increasing frequency in advanced microlithography. The degree of polarization is degraded when the light passes through a mask substrate with a substantial level of birefringence. Fortunately, stress birefringence can be minimized by careful thermal control of the substrate manufacturing processes. Because this adds to the manufacturing costs, a low birefringence requirement is specified only for masks that will be used with polarized illumination.

5.2.3 Requirements for absorber films

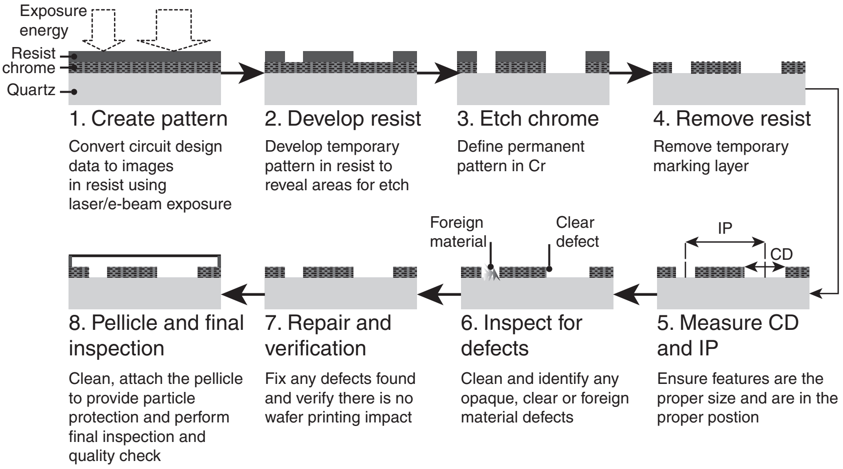
The light-absorbing film used to make the mask pattern must be free of embedded particle defects or pinholes larger than a few nanometers. Finished mask blanks are usually tested by an optical inspection system that scans the film surface at high resolution, using transmitted light to look for pinholes and reflected light to look for embedded particles.

Although the absorber film is very thin, the sputter deposition methods used to make the film may leave it with high levels of either tensile or compressive stress. This can induce a bowed mask surface that may exceed the stringent flatness requirements for a finished mask. When the mask pattern is etched into the absorber film, some of the film stress will be relaxed, but in a complex way depending on the local pattern density in different areas of the mask. Because of these undesirable effects, the mask blank manufacturers try to tune the deposition process to minimize stress in the absorber.

The front surface of the absorber film must not be too reflective at the exposure wavelength, or scattered light from the projection lens, the lens mounts, or the wafer surface will be reflected into the image, reducing the contrast of the image or even producing ghost images. A smooth chromium surface has a reflectivity of 45–65% at commonly used lithographic wavelengths. This can be reduced to a reflectivity of 10–20% by use of a thin, anti-reflective layer of chromium oxide or chromium oxynitride deposited on the top surface of the chromium. Other mask absorbers materials (such as molybdenum silicide) are also deposited with one or more surface layers to control reflectivity.

There is a different set of requirements for optical properties at the wavelengths used to inspect the masks for defects. The contrast between the absorber and the fused silica substrate must be sufficient to distinguish the mask features. There can also be problems with detector saturation in the inspection equipment if the absorber reflectivity is too high. The conflict between different optical properties required for the exposure wavelength and the inspection wavelength can make the optical design of the absorber film quite challenging for the blank supplier.

The absorber film should be as thin as possible. Most lithographic simulation programs use a thin-mask approximation (TMA), which assumes that the mask patterns are purely two dimensional structures with no physical thickness. Because of light interaction with the vertical sidewalls of an actual mask pattern, the simulations become increasingly inaccurate with increasing pattern thickness. Lithographic simulation programs are extensively used in mask design to correct for non-linear imaging behavior in the projection optics, and it is extremely important that the simulations match physical reality as closely as possible. Simulation programs exist that can model the full electromagnetic field (EMF) interaction between light and a three-dimensional absorber structure, but these programs run several orders of magnitude slower than the TMA simulations, and are not practical for full-scale mask designs. So far it has been more practical to invent thinner absorber films than to use simulation programs that include all the optical complexities of thick absorber films. Typical absorber films range between


Exposure energy
Resist
chrome
Quartz
1. Create pattern
Convert circuit design data to images in resist using laser/e-beam exposure
2. Develop resist
Develop temporary pattern in resist to reveal areas for etch
3. Etch chrome
Define permanent pattern in Cr
4. Remove resist
Remove temporary marking layer
Foreign material
Clear defect
IP
CD
8. Pellicle and final inspection
Clean, attach the pellicle to provide particle protection and perform final inspection and quality check
7. Repair and verification
Fix any defects found and verify there is no wafer printing impact
6. Inspect for defects
Clean and identify any opaque, clear or foreign material defects
5. Measure CD and IP
Ensure features are the proper size and are in the proper postion

5.2 Mask process flow from pattern formation through pellicle mount.

40 and 100 nm thick, with the films at the upper end of the range exhibiting the worst agreement with thin-mask simulations.

There is a final requirement for a mask film, unrelated to its optical properties. The film must have good electrical conductivity to prevent build-up of electrical charge during the electron beam mask writing process or scanning electron microscope (SEM) measurements. Materials in current use have acceptable conductivity, but a film with very low conductivity would require a conductive top layer to enable accurate mask patterning with an electron beam.

5.3 Mask process

In the 1970s, masks were made by carving a colored film by hand and using multiple reductions to create the final mask for wafer exposure. Today, masks are fabricated in a clean room facility known as a mask house, mask shop, or mask fabricator. A specific mask process description is offered as an example, because providing exhaustive descriptions of all mask permutations is not practical. One of the simplest process flows is based on a COG binary mask, and this is what is shown in Fig. 5.2. Descriptions of the steps are contained within this chapter: processing steps are in Section 5.3, metrology in Section 5.4, and work related to finding and controlling defects in Section 5.5.

The starting mask blank includes the glass substrate and the unpatterned absorber layer. Many mask fabricators receive the blanks with photoresist precoated, but this practice risks expiration of the resist film. Before being coated, the blanks are inspected for incoming quality and cleaned to remove any organic contamination on the chromium surface. To enhance photoresist adhesion to the chromium, the mask is baked at a low temperature, say 120°C, and a molecular monolayer of hexamethyldisilazane (HMDS) is added to promote adhesion. Photoresist dissolved in an organic solvent is delivered to the mask surface as the mask is spun at a few thousand revolutions per minute to distribute the film evenly in a process called spin coating. The mask is then baked again to remove solvents from the resist.

The resist is typically patterned with either laser or electron exposure. Electron beams, known as e-beams, are appropriate for high-end applications because the electrons can be focused and deflected with electromagnetic or electrostatic elements to form a nanometer-scale spot. The e-beam systems can be designed to raster across the entire mask, turning off and on as needed, or to be vector scanned only in the desired regions. Accelerating voltages are in the range of tens of keV, and current densities are 70–200 A/cm^2. Laser exposure is generally faster, but cannot achieve the same performance as e-beams because the spot size is limited by optical diffraction. The laser used is in the UV region ~250–450 nm and multiple beams are raster scanned. The purpose of both laser and e-beam exposure is to change the chemistry of the photoresist. The resist bonds are either broken down to dissolve in the developer (positive resist), or are cross-linked to become insoluble (negative resist). The resist is developed, either by tank immersion in the developer or, for better defect performance, by a spin/puddle process. A common developer is an aqueous solution of tetramethylammonium hydroxide (TMAH).

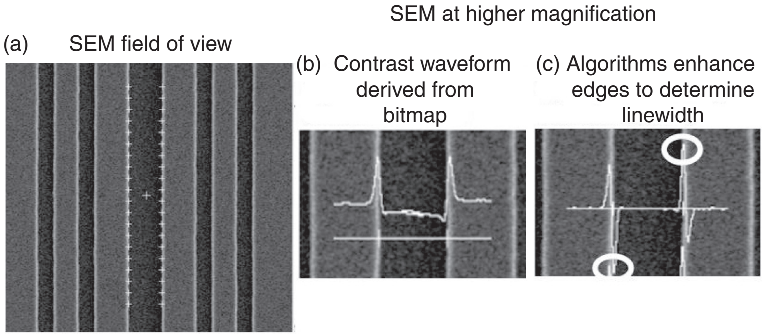
A light oxygen descum process can remove artifacts of an incomplete develop at the bottom of developed mask features. A chromium etch then removes the absorber material selectively. This is called pattern transfer, because the resist pattern is transferred into the absorber layer. For chromium absorbers, a wet etch process using perchloric acid (or other acids) and ceric ammonium nitrate has been used. The isotropic nature of the wet process limits performance because of the unintended lateral etching when removing material. Anisotropic dry or plasma etching of the absorber is more common for applications requiring good lateral control of features. Plasmas with Cl_2 and O_2 are typically used for chromium etching. Selectivity of the absorber etch to the patterned photoresist material is important because a thinner resist layer can be used to protect the unetched areas, and thinner photoresist improves resolution when other factors are held constant.

After pattern transfer, the remaining photoresist film is stripped. The strip can be accomplished with wet chemistry, such as a sulfuric acid and hydrogen peroxide mixture (SPM), to remove photoresist and organics. It can also be removed with an oxygen-containing plasma. A final clean is required to remove any particles and adsorbed surface contamination. A common chemical mixture for removing organic material and particles in the semiconductor industry is known as “Semiconductor Clean 1” or SC-1. This clean is commonly applied to masks as well as wafers. The chemistry is primarily water with small amounts of hydrogen peroxide and ammonium hydroxide; a typical example of dilution ratios is 80:2:1. Megasonic energy is added to enhance particle removal efficiency.

After the final clean, the binary mask is complete from an image formation perspective. The photomask qualification demands at least half of the process time and has not yet been described. It includes critical dimension (CD) metrology, image placement (IP) metrology, inspection for feature integrity, and repair of any defects. It is the stringent requirements in these areas that make photomask manufacturing difficult. Another obstacle is that each photomask design is essentially unique, a custom-made component in the microlithography process.

5.4 Mask metrology

Metrology is a general term for the science of measurement. Mask metrology is used to verify the quality of the mask pattern by measuring a selected



SEM at higher magnification

(a) SEM field of view

(b) Contrast waveform derived from bitmap

(c) Algorithms enhance edges to determine linewidth

5.3 CDSEM images showing (a) the normal field of view, (b) the extracted waveform, and (c) an algorithm is applied to the waveform to enhance edges, circled for emphasis.

group of features across the mask. This metrology sample is selected to be representative of the overall mask. The number of features included in the metrology sample must be large enough to ensure that the mask features meet specifications across the mask, but not so large that the metrology time is unreasonably long. The primary metrics that are measured on a photomask are: the feature dimension, the placement of the feature relative to an absolute grid, and the light modification properties. The corresponding metrology metrics are known as the CD, the IP, and phase and transmission.

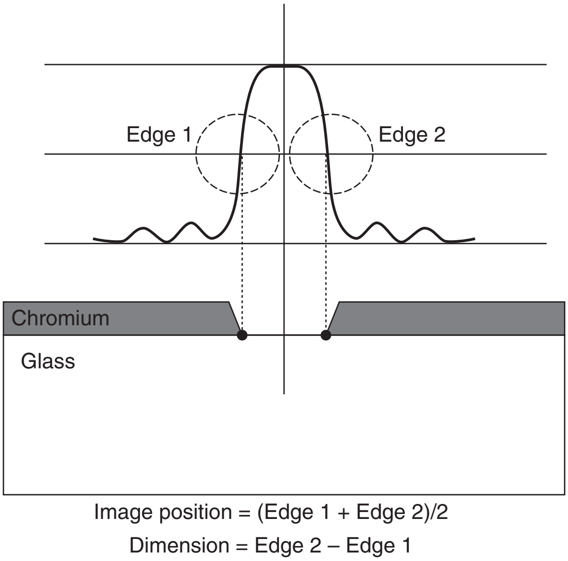
5.4.1 Critical dimension (CD) metrology

The feature size, or CD, is measured with an optical microscope, an SEM, or an atomic force microscope (AFM). Historically, only a simple one-dimensional width of the feature was measured. As lithographic fidelity has improved, secondary features of the mask may transfer to the wafer so that line edge roughness, corner rounding, and general 2D shape fidelity have become important. Optical microscopes have insufficient resolution for many applications. AFMs have high resolution, but are very slow. SEMs offer high resolution and reasonable speed, and dominate mask CD metrology applications. SEMs optimized for top-down CD measurements of semiconductor patterns or masks are often called CDSEMs. There are several steps to the image acquisition and processing in a CDSEM. The electron beam is scanned across the mask, so that the mask surface emits secondary electrons. Three stages of CDSEM data extraction are shown in Fig. 5.3. First, the emitted electrons are recorded by detectors to create a two-dimensional image. Then, the contrast waveform is extracted from the bitmap and the edges of the features become visible as intensity peaks. Finally, the edges, are sharpened by applying algorithms so that the CD is given by the distance between the two edges. A known pitch standard is often used to calibrate the SEM output to a physical dimension. It is important to acknowledge that the underlying image is an approximation of reality; calibrating the results to standards is essential for accurate CD results.

5.4.2 Image placement metrology

As pattern sizes shrink through the continuing evolution of semiconductor technology, the requirements for accurate placement of mask features are also becoming tighter. For the most advanced semiconductor designs, pattern placement accuracy requirements are now in the 5–10 nm range. Because the mask is magnified 4× to 5× the size of the chip pattern, one might imagine that the mask IP tolerance would be correspondingly magnified; however, the mask contribution is only one small part of the total IP budget on the wafer. Consequently, the mask IP tolerances are also in the 5–10 nm range.

After a mask pattern has been created, IP is measured by an extremely accurate optical measuring device. The finished mask is placed on an interferometer-controlled stage, and an array of registration marks is viewed by a scanning laser beam. In order to support the required mask tolerances, subnanometer repeatability is required for the IP measurement. The two edges of the feature are identified and the location, or placement, of the image


Edge 1
Edge 2
Chromium
Glass
Image position = (Edge 1 + Edge 2)/2
Dimension = Edge 2 – Edge 1

5.4 A schematic diagram showing the optical signal from a clear feature on a mask and the calculations for CD and IP.

is calculated as shown in Fig. 5.4. The image size can also be extracted, but since the measurement is optical, this ancillary measurement may be insufficiently accurate for CD metrology requirements.

A large number of registration marks, typically several dozen to a few hundred, will be measured on a photomask. This allows increased accuracy by averaging many measurements and provides a detailed view of localized mask distortions. The raw map of IP measurements is analyzed to remove errors with low spatial orders, such as whole-body displacement, rotation, magnification, and orthogonality. All of these low-order terms can be corrected by the wafer exposure equipment. The higher-order errors cannot be easily corrected, and limit the ultimate layer-to-layer overlay performance that can be achieved with the mask. Because of the extreme accuracy requirements of the IP measurement, the temperature and humidity environment inside the metrology equipment are tightly controlled. Even the small gravitational sag distortion of the mask induced when the mask is held horizontally during the measurement process is mathematically corrected to achieve the final sub-nanometer repeatability of the measurements.

5.4.3 Phase and transmission metrology

Photomasks that use a phase-shifting attenuator film require additional measurements to ensure that the optical phase shift and transmission of the film are within tolerances of their design values. Because the phase shift and transmission are both wavelength dependent, it is necessary to do the measurement with a light source having the same wavelength used to expose the pattern on the wafer. The same type of excimer laser used to expose the wafer could be used for phase and transmission metrology, but it is easier to use a broadband deuterium plasma lamp or mercury–xenon arc lamp with a monochromator to isolate the desired wavelength.

Once a beam of monochromatic light is produced, conventional dosimetric methods can be used to measure the film transmission. Two different conventions are used to normalize the transmission measurement. Either the light transmission through the mask substrate and absorber film is measured relative to an equivalent air path, or the transmission through the film-covered region is normalized to the transmission through a clear region of the fused silica mask substrate. Although fused silica absorbs very little light at commonly used lithographic wavelengths, there are substantial reflective losses from the surface of the mask. For purposes of lithographic simulation programs, the transmission relative to a clear area of the mask is more useful than transmission relative to an air gap.

Optical phase metrology is more difficult than transmission metrology, and requires some sort of interferometric measurement. The most practical arrangement has been to use a shearing interferometer, which produces interference fringes by splitting the image of a small mask pattern and sending the two images down different light paths before recombining the two images with a small transverse offset between them. By analyzing the spatial offset between the interference fringes in different parts of the recombined image, the optical phase shift of the mask film can be derived. Commercial instruments are available that can automatically measure phase shift and transmission in patterned phase-shifting attenuator films.

5.5 Defects and masks

Defects are called soft if they can be removed by cleaning, and hard if they cannot. Soft defects include particles, stains, and contamination. Hard defects include extra or missing absorber, scratches, and film quality problems such as bubbles. Either type of defect is a risk to the quality of the printed image, because the mask is the master template for the subsequent wafer patterning. A single mask can be replicated 100 times on each of 1000 wafers when fabricating a batch of semiconductor chips. A single error on the mask has the potential to impact over 100 000 devices. This risk is clearly unacceptable; so mask inspection, repair, and verification have become the cornerstone of mask quality assurance.

5.5.1 Patterned mask inspection

Mask inspection is not generally considered a metrology method, since the process does not involve measurements of the mask features. The primary purpose of the inspection is to scan the mask for defects. The process scans every pattern on the mask instead of relying on a limited number of features to represent the entire mask. Consequently, it is an integral part of final quality, process learning, and defect density understanding. It is possible to re-analyze the complete mask scan to create CD and IP maps, but this remains a secondary function of the inspection tools. The majority of patterned mask inspections are performed with bright-field optical microscopes. Ideally, the wavelength of the inspection system matches the exposure wavelength, so that any wavelength-dependent impact is captured. More typically, the wavelength lags the technology node. For example, the bulk of 193 nm lithographic masks are inspected using 257 nm light. Both transmitted-light and reflected-light inspection modes are in use.

There are two options for patterned inspection modes. Die-to-die inspection is possible only when there are multiple repeating fields, or dies on the mask. One area of the mask is compared to an identical region on the same mask, and the differences are identified as defects. The most widely applicable inspection mode is die-to-database, which compares the mask to a database version of the mask-design data that have been modified to capture process distortions. A third type of inspection is possible only on optically transparent masks (not EUV masks, for example). It combines both transmitted-and reflected-light signals to create a signal. The signal is calibrated on representative features and then the full mask is scanned. Any signal that does not cleanly match the combined transmission and reflection signature of the calibrated mask structures (i.e. clear areas and absorber areas) is identified as a defect. This type of inspection is effective for detecting transmission defects like stains or glass damage and reflectivity defects such as particles, but is not appropriate for small feature defects. It requires neither a matching die for pattern comparison, nor a database definition of the pattern; so it is a general-purpose method for detecting these types of defects.

5.5.2 Mask repair

If a defect in the mask pattern is discovered there are only two options: either the mask must be discarded and rebuilt, or the defect must be repaired. Because of the high cost of manufacturing a photomask, the better option is to repair the defect if possible. In the very earliest days of lithographic mask making, this repair sometimes consisted of daubing spots of ink on the mask to replace missing bits of absorber, while holding one's breath and peering through a microscope. As mask feature sizes descended over time into the micron and sub-micron range, this method rapidly became impractical.

The next advance in mask-repair technology was the invention of a laser repair method. Short-pulse lasers, typically solid-state yttrium aluminum garnet (YAG) lasers, were developed to burn away unwanted parts of the chromium pattern. The laser beam could be focused down to a sub-micron spot size, and laser pulses in the 10 ns range could raise the temperature of chromium within the focus spot high enough to vaporize it. There were several problems with this repair method, but for many years it was the only method available for removing opaque pattern defects. If a large amount of chromium was being removed, the vaporized material would often condense and be redeposited around the area of the repair. This could be controlled to some extent by flushing the area around the repair with a fast air flow, to sweep away the debris before it could condense. The laser pulse was aimed manually, by an operator viewing the repair site through a microscope, and the placement of the repair was extremely difficult to control. The thermal conductivity of chromium is very high compared to fused silica; so the heat of the laser pulse was largely confined to the chromium layer, but the heat would spread horizontally through the chromium film, producing a region of melted chromium around the edges of the laser spot. This melted chromium would flow a short distance and form a thick bead from surface tension. If the edge of the cleared zone was not in the correct location, it was usually impossible to vaporize the thickened edge bead without causing more damage to the adjacent thin chromium.

Laser repair could only remove material; it could not be used to repair clear defects. Two methods were developed for clear defect repair, using optical equipment similar to the laser repair equipment. In the first method, the mask was coated with photoresist and a beam of white or ultraviolet light was focused onto the photoresist above the clear defect. After the exposed area had been removed by a chemical developer, a new layer of chromium was sputtered onto the surface of the mask. Then the remaining photoresist was dissolved in a solvent, allowing the new layer of chromium to wash off everywhere except in the small region over the defect, where it had bonded to the mask. This method, called the chrome lift-off method, was complex and slow, and had the additional problem of creating clouds of chromium debris that had to be washed from the mask surface. The second method of repairing clear defects used a focused laser beam to thermally decompose a vapor such as chromium hexacarbonyl, which would create a thick, chromium-rich plug over the defect that was being repaired. This method, like chrome lift-off, had severe drawbacks. The material being deposited had a ragged edge profile, rather like an arc-welder's bead. The deposition process could be started only by focusing the laser on a chromium region; if the laser was first focused on a clear area of fused silica, then not enough heat would be generated to start the decomposition of the hexacarbonyl vapor. Still, the method worked well for repair of isolated pinholes in the chromium surface, and could be used for repairs on the edges of features if the edge quality was not critical. All of these methods required highly skilled operators, who placed the repairs visually.

The next generation of mask-repair technology began with the invention of focused ion beams. Although ion implantation has been used in semiconductor manufacturing since the 1970s, high-brightness liquid metal ion sources were developed only in the mid 1980s (Swanson, 1983). Ions, typically gallium ions, emitted from these sources can be focused to a sub-micron spot, and have enough energy to remove material from a mask surface by sputtering. By scanning the ion beam across the mask surface like an SEM, an image of the mask surface can be captured. By allowing the ion beam to dwell on areas of the absorber pattern that need to be removed, these areas can then be sputtered away. A method of repairing clear defects was also developed, using the focused ion beam to decompose organic vapors such as styrene at the mask's surface (Harriott, 1988). The graphitic material resulting from the decomposition adheres strongly to the mask surface and makes a durable patch. Because the scanning ion beam is able to detect an image of the mask pattern, signal processing technology allows the repairs to be automatically aligned to the edges of existing mask features, enabling repairs that are no longer dependent on operator skills.

Focused ion beam technology brought a marked improvement to the quality of mask repairs, but it still has a few negative qualities. The energetic gallium ions used to sputter away the absorber defect become implanted in the transparent mask surface, reducing the transmission by a substantial amount. Although the loss of transmission can be partially compensated by distorting the mask shapes, enlarging the size of clear features and reducing the opaque features, this solution was not always satisfactory. A large reduction in gallium staining was achieved by adding a reactive gas to the region around the repair site. The gas is activated by the ion beam and removes the defective parts of the mask pattern by chemical etching, greatly reducing the ion dose required. It should be noted that these chemical etching processes are very material dependent, and a process designed to remove a chromium defect may have no ability to remove a stray hydrocarbon-based stain from residual photoresist.

There have also been advances in laser repair technology. By reducing laser pulse lengths from 10 ns to the 100 fs range, the problem of chromium film melting has been eliminated. The extremely short pulses vaporize the surface of the absorber film before the heat can be conducted away from the beam spot, giving a very clean edge to the repair. Automated alignment is also used in newer laser repair tools, greatly improving alignment accuracy of the repairs.

Micromachining technology has been adapted for use in mask repair. The microscopic probes used for AFM can be made strong enough to physically scrape away unwanted bits of the mask pattern. The same probe tip used for the repair is also used to capture profiles of the mask before and after the repair. The great advantage of this repair method is its insensitivity to the nature of the mask defect. Surface particles and embedded defects in the mask film are removed with the same process used to repair pattern defects.

Electron beams can also be used for mask repair. Although electrons do not have enough mass to be useful for material sputtering, the gas-assisted etch and deposition processes developed for ion beam repair can be deployed with electron beams. The electron beam does not produce the staining seen with implanted metal ions, making the process much cleaner. Advanced electron beam repair equipment can reconstruct missing areas of the mask pattern using the original electronic design data as input, and automatically aligning the reconstructed patterns to the surrounding areas of the mask.

5.5.3 Pellicles

After a mask is fully repaired and cleaned, the images it produces on a semiconductor wafer are essentially free of defects. Unfortunately, this state of affairs is only temporary. No matter how clean the environment around the mask is, a few stray dust particles will eventually be deposited on the mask. A structure called a pellicle is used to prevent these particles from landing on the critical patterned surface of the mask.

A pellicle is a thin membrane made of an organic or fluorocarbon polymer. It is stretched tightly over a stiff, flat frame, and the frame is glued to the mask surface surrounding the critical patterned area. The pellicle is extremely thin, on the order of one micron or less. The spacing between the pellicle and the mask pattern depends on the height of the frame, and is typically a few millimeters. Any dust particles that land on the surface of the pellicle are suspended far from the patterned surface, and are so far away from the focal plane of the projection optics that they do not appear in the projected image. Of course, a large enough dust particle on the pellicle will produce a shadow in the image plane, but the sensitivity to particle size on the pellicle is many orders of magnitude less than it is on the mask surface.

The pellicle and frame protect the mask from particles, but they also create a volume of trapped air between the mask and pellicle. Changes in barometric pressure or changes in altitude experienced during air shipment of a mask with pellicle can inflate the pellicle surface enough that it may contact parts of the shipping container, mask library, or optical exposure equipment. To prevent this, ventilation holes are provided in the pellicle frame. Small particle filters are placed over the vent holes to prevent particles from being pulled in with the air flow.

The pellicle membrane is so thin that it exhibits strong thin-film optical interference effects. For any given film thickness, the optical transmission will vary with angle of incidence and wavelength. Because the mask is exposed with monochromatic light, the thickness of the pellicle can be tuned to give a maximum transmission greater than 99% at the exact exposure wavelength. The thickness variations across the pellicle surface must be small enough to keep the transmission above 99% everywhere. Because of the angular dependence of the pellicle transmission, there will be some light loss for mask illumination that is not at normal incidence. Modern lithographic projection optics can illuminate masks at angles of incidence up to 20°. A one-micron thick pellicle will have a 9% or greater transmission loss at this angle. To keep the transmission loss to a minimum when angles of incidence are this high, a much thinner pellicle must be used. A pellicle less than 0.3 microns thick can have greater than 99% transmission over a range of incident angles from 0° to 20°. Pellicles for use at 365 nm or longer wavelengths can be made with thin, anti-reflective layers, which greatly reduce the sensitivity to wavelength and angle of incidence. Unfortunately, no acceptable materials are known that can be used for antireflection coatings on pellicles used for 248 nm or shorter wavelengths.

The pellicle sits in the optical path between the mask pattern and the wafer, and effectively becomes a part of the projection optics. This raises questions about the effects of pellicle motion on the final projected image. Although the pellicle membrane is anchored firmly to the pellicle frame, the unsupported center areas of the pellicle can flutter by several tens of microns in ambient air currents. Optical analysis shows that small vertical displacements or tilts of the pellicle membrane have vanishingly small effects on the image quality, and that these types of pellicle motions should not cause concern. However, relatively small gradients in the pellicle thickness can induce optical aberrations in the images. A linear change in pellicle thickness over a short horizontal range will induce a tilt aberration that affects IP, and more complex thickness changes will induce higher-order aberrations in the images. Although thickness gradients in a pellicle are rarely seen, they can be caused by pellicle manufacturing errors or degradation of the pellicle membrane from optical exposure. The effects of thickness gradients in pellicles have not been widely discussed in the lithographic industry, but they remain a potential problem.

Pellicles membranes, frames, and the adhesives used to secure the pellicle to the frame, and the frame to the mask, are all potential sources of contamination to the mask. The aluminum frame is often built with a rough, black anodized surface to reduce reflections, and has a potential for shedding particles inside the protected area of the mask. The other components are made of polymeric compounds and may contain non-negligible amounts of residual solvents. Organic compounds that outgas from the pellicle membrane or adhesives can interact with UV light and the mask surface to produce small condensed particles that grow with continued exposure until the mask is no longer usable. The development of low-outgassing pellicle adhesives is an active area of development.

The mask industry has become comfortable with qualifying the mask, protecting it with a pellicle and using it with confidence that no particles will be added during exposure. Mask inspection in the wafer fabricator does exist, but is typically used for quality insurance and not as a daily method for ensuring that no new particles have settled on the pattern areas. For all the mask types introduced in Table 5.1, except the optical ones, there is no pellicle option. The absence of a pellicle forges an even stronger connection between microlithography and mask processing, since the masks will need to be cleaned and inspected frequently, and subtle changes must be monitored and understood. The mask maker's job is to ensure that wafer yields are not impacted by mask changes. This mission underscores all of the mask work outlined in this chapter.

5.6 Conclusion

This chapter has reviewed the general concept of a mask as it is used in microlithography. The concept of replicating a two-dimensional design is simple. Fabricating functional photomasks to be used in microlithography can be challenging for many reasons: material selection, resolution requirements, sensitivity to defects, and the fact that the tolerances for IP and CD are so small. Some details of the mask-making process and quality measurements were introduced. Length limitations necessitated only a brief treatment of complex topics in this general chapter on photomasks. For deeper treatments, please consult the references, in particular the two photomask books, one edited by Eynon and Wu (2005) and the other by Rizvi (Rizvi, 2005). The International Technology Roadmap for Semiconductors (ITRS, 2011) contains a very useful summary of today's mask specifications and an industry consensus of future mask requirements.

5.7 References

Anon (1982), *Optical Crystals Handbook 82*, Optovac, Inc., North Brookfield, Massachusetts.

Batchelder D N and Simmons R O (1964), "Lattice constants and thermal expansivities of silicon and of calcium fluoride between 6° and 322° K", *Journal of Chemical Physics*, **41**, 2324–2329.

Corning (1991), *Fused Silica, Technical Products Division*, Corning, Inc., Corning, New York.

Corning (2006), *ULE Corning Code 7972, Ultra Low Expansion Glass, Corning Advanced Optics*, Corning, Inc, Corning, New York.

Davila L, Risbud S and Shackelford J (2010), 'Quartz and silicas', in Shackelford J and Doremus R (Eds), *Ceramic and Glass Materials: Structure, Properties and Processing*, Springer-Verlag, New York, 75–86.

Dickinson S (1975), "Infrared laser window material property data for ZnSe, KCl, NaCl, CaF2, and BaF2", *AFCRL-TR-75–0318, Air Force Cambridge Research Laboratory Technical Report*.

Eynon B and Wu B (2005), *Photomask Fabrication Technology*, McGraw-Hill, New York.

Harriott L R and Vasile M J (1988), "Focused ion beam induced deposition of opaque carbon films", *Journal of Vacuum Science and Technology B*, **6**(3), 1035–1038.

Hosono H, Mizuguchi M, Skuja L and Ogawa T, (1991), "Fluorine-doped SiO2 glasses for F_2 excimer laser optics: fluorine content and color-center formation", *Optics Letters*, **24**(22), 1549–1551.

ITRS, International Technology Roadmap for Semiconductors. Available from: www.itrs.net.

Rizvi S (2005), *Handbook of Photomask Manufacturing Technology*, Boca Raton, Taylor & Francis.

Sze M (1983), *VLSI Technology*, McGraw-Hill, New York.

Swanson L W (1983), 'Liquid metal ion sources, mechanism and applications', *Nuclear Instruments and Methods*, **218**, 47–353.

Yamagata S, Inaki K, Matsuya T, Takke R, Thomas S and Fabian H (Shin-Etsu Quartz Products Co, Ltd. and Heraeus Quarzglas GmbH) (1992), *Optical members and blanks of synthetic silica glass and method for their production.* US patent 5325230, 4 February 1992.

6

Maskless photolithography

M. E. WALSH and F. ZHANG, LumArray, Inc., USA,
R. MENON, University of Utah, USA, and H. I. SMITH,
LumArray, Inc. and Massachusetts Institute of Technology, USA

DOI: 10.1533/9780857098757.179

Abstract: The virtues of using photons for maskless lithography rather than charged particles are discussed briefly. Two generic forms of maskless photolithography, image projection and focal-spot writing, are compared. A particular form of focal-spot writing, zone-plate-array lithography (ZPAL), is highlighted. In this approach, an array of diffractive-optical microlenses is employed to produce stationary, discrete focal spots. Patterns are created by modulating the focal-spot intensities under computer control as the substrate is scanned. The reasons behind the design choices in ZPAL, experimental results, commercialization efforts, and prospects for circumventing the diffraction barrier are described.

Key words: maskless lithography, maskless photolithography, zone-plate-array lithography, focal-spot writing, diffractive-optical lenses, absorbance-modulation optical lithography.

6.1 Introduction

6.1.1 Mask-based and maskless lithography

With few exceptions, structural complexity, and hence functionality, is produced in electronic, photonic, mechanical and mixed systems via the planar process, consisting of lithography followed by some form of pattern transfer. Lithography is the component that puts ‘information’ into the pattern. Among lithographic systems we distinguish between those that replicate a preexisting structure (a mask or template) and those that create a pattern, so-called maskless-lithography systems, sometimes called pattern generators.

6.1.2 The printing press and the desktop printer

Lithographic systems that use a mask or templates are analogous to a printing press for document production, in that a master is made and then

Published by Woodhead Publishing Limited, 2014

reproduced on the press. The limitations of master-based printing are obvious, when compared to using word-processing software and printing via a direct electronic connection. The latter are more suitable when documents are prepared in limited numbers, when modifications to a document are desirable, or when multiple small variations on a given document are needed. From the viewpoints of convenience, flexibility, innovation, and productivity, maskless lithography is analogous to document production via a computer-driven printer. For this reason, maskless lithography is clearly appropriate for research and development, and also for customized and low-volume manufacturing, provided the throughput and quality are sufficient. It is highly probable that efficient, moderate-throughput maskless-lithography systems, once fully developed, will make customized manufacturing more attractive, profitable, and widespread.

A variety of maskless-lithography systems have been developed, some using electrons, some using ions, and others using photons. Each approach has its advantages and disadvantages. It is instructive to discuss these before narrowing the focus of this chapter to maskless photolithography, or a particular mode thereof.

6.2 The use of photons as opposed to charged particles

6.2.1 Instability of charged-particle systems

Scanning-electron-beam lithography (SEBL) has been available for over five decades and is widely used in research and, to a limited extent, in low-volume manufacturing. Since electrons and ions can be focused down to spot sizes of the order of a nanometer, it is relatively easy to create patterns with deep-sub-100 nm resolution. This is the major advantage of charged-particle lithography systems. However, SEBL and other forms of charged-particle lithography, such as shaped-beam projection, have a number of significant problems. Low throughput, resulting from the mutual repulsion of charged particles, is one that is well known.[1,2] Ideas for circumventing this problem, such as a massively-parallel array of independently controlled beams, have been proposed.[2–4] However, a significant problem that is usually overlooked is the basic instability of charged-particle systems, as manifested in pattern-placement errors. The position of an electron beam can drift relative to a substrate as a result of charging, stray electric or magnetic fields, and thermal gradients. A feedback approach to solving this problem for a single-beam system has been proposed and demonstrated[5,6] but implementing it for thousands of beams is probably not feasible. In a multibeam system, reducing the interaction among beams to negligible levels, as they switch on and off, may prove highly problematic.

6.2.2 Benefits of photons

In contrast to charged particles, photons are not deflected by electric or magnetic fields. They do not interact with one another and hence no limits are imposed on the number of photons that can be concentrated in a given space. The photons used in near-UV lithography have low energies (~3eV) and produce minimal damage to substrates, including organic substrates. Shot-noise effects are negligible compared to those exhibited by high-energy particles.[7] The phase and polarization of light can be controlled and used to advantage in imaging. Photons retain their frequency and phase until absorbed, whereas charged particles lose energy and phase information as they traverse material. Lasers provide beams of very high brightness.

Photons in the visible and near-UV bands interact with outer electrons. As a result, one can invoke wavelength-selective chemistry to achieve a wide range of lithographic advantages. Antireflection (AR) coatings are an example. When placed on top of a photoresist film, an AR coating acts as an impedance match between air and photoresist, increasing the coupling of light into the film and reducing standing-wave effects. If the top AR coating is a bleachable absorber, it can enhance image contrast by suppressing low-level radiation relative to high intensities. When placed underneath a photoresist film, an AR coating serves to eliminate standing waves by strongly absorbing light, preventing it from reflecting back into the photoresist. In optical microscopy, wavelength-selective chemistry has been used to achieve resolution down to 1/40th of the optical wavelength.[8] In Section 6.7 we discuss prospects for achieving similar resolution in photolithography.

6.3 Forms of maskless photolithography

6.3.1 Two generic forms: image projection and focal-spot writing

We distinguish two generic forms of maskless photolithography: image projection and focal-spot writing. In image projection, a two-dimensional spatial-light modulator, containing a large number of picture elements (pixels), is used in place of a mask. This is the approach taken, for example, by Micronic Laser Systems AB (http://www.micronic.se/). Image projection generally requires that the stage holding the substrate be stationary, as the image is exposed in resist. Because the 'spatial bandwidth' of any projection lens is limited, the number of pixels in the projected image is likewise limited. Expressed another way, there is a direct coupling between the image resolution and the image size, a coupling that is familiar to anyone who has used an optical microscope. Moreover, every projection lens has some aberration and distortion, and although this can be small, and is repeatable, for some applications such imperfections can be problematic.

Partially-coherent illumination is generally used in image projection, which complicates certain aspects of lithography. For example, as one approaches image-resolution limits, the pattern on the mask deviates significantly from the image in resist. If the desired pattern on the substrate is a rectangle, serifs must be added to the ends of the rectangle on the mask; otherwise the rectangle image would be narrower and shorter than desired. In general, correction for the effect of nearby patterns (proximity-effect correction) is extremely challenging with partially coherent imaging.[9]

6.3.2 Maskless photolithography via focal-spot writing

In focal-spot writing, one or more beams are focused on a substrate and moved across it as intensities are modulated so as to create a desired pattern. In effect, patterns are created by overlapping focal spots. Several distinct approaches are possible. One can: (1) use a single lens, focus a single beam through it, and scan the substrate; (2) use a single lens, focus a single beam through it, and scan both the beam and the substrate (http://www.himt.de/en/hom); (3) use a single lens, focus several beams through it, and scan both the beams and the substrate;[10] or (4) use several lenses, focus single beams through each, and scan only the substrate.[11,12] In these focal-spot-based approaches there is no coherent relationship between successively overlapped focal spots, and hence there are no artifacts due to interference, as occur with coherent or partially-coherent imaging systems. Focal-spot intensities are additive, that is, linear superposition applies, and the imaging is 'incoherent.' This greatly simplifies the lithography, especially proximity-effect correction and linewidth control.

The resolution limit, defined as the minimum resolvable grating period, $p_{\min}$, for incoherent imaging is given by

$$P_{\min} = \frac{\lambda}{2\sin\theta} = \frac{\lambda}{2\mathrm{NA}} \qquad [6.1]$$

where λ is the wavelength, θ is the half angle of the light converging to the focus, and NA is the 'numerical aperture' (NA). This implies that there is a small resolution advantage for incoherent imaging relative to partially-coherent imaging. But the more important advantage of incoherence is the absence of interference effects and the validity of linear superposition.

Because focal-spot writing can extend over an entire substrate, or any fraction thereof, there is no coupling between the resolution of the written pattern and its spatial extent. This decoupling of resolution and area greatly expands the efficacy and applications of maskless photolithography, and for this reason the rest of this chapter will discuss only focal-spot writing, concentrating on the multi-lens approach (#4 above) since this has potential for the highest throughput, the highest precision and accuracy, deep-sub-100 nm resolution, and is the most recently introduced. Also, a commercial system is under development that will make the technology available to the scientific and engineering communities.

6.3.3 Focal-spot writing with an array of lenses

The throughput of focal-spot writing is directly proportional to the number of independently modulated focal spots. This is expressed by Equation [6.2]:

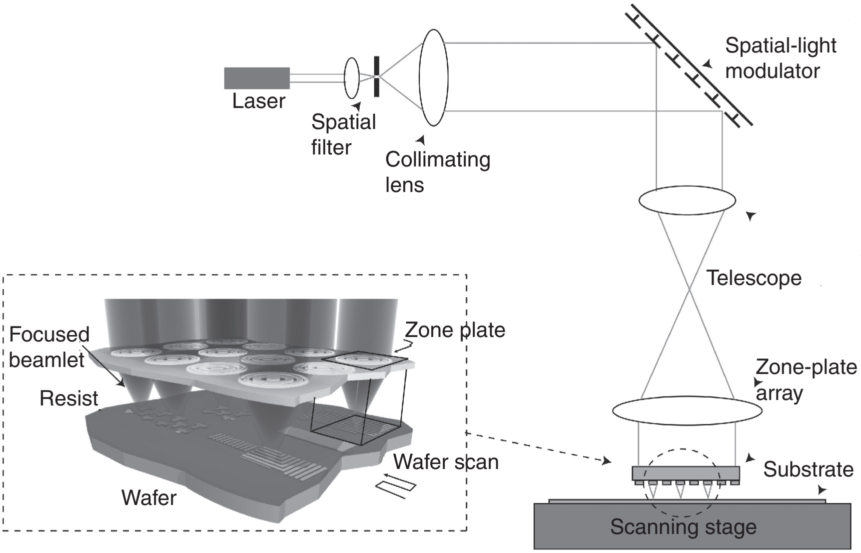
$$A = NRd^2 \tag{6.2}$$

where A is the exposure rate in area per unit time, N is the number of independently modulated focal spots, R is the update rate of the modulator, and d is the exposure-grid spacing. To achieve smooth line edges, the exposure-grid spacing, d, should be half the focal-spot diameter, or finer.

6.4 Zone-plate-array lithography (ZPAL)

While it is feasible to carry out multi-lens focal-spot writing using an array of refractive lenses, it is preferable to use an array of diffractive-optical lenses for a variety of reasons discussed below. This form of focal-spot writing is called *ZPAL*. Initial studies used an array of binary-phase Fresnel zone plates.[13–16] Figure 6.1 illustrates schematically the basic elements of a ZPAL system: a source of collimated, continuous radiation; a spatial-light modulator; optics that interface the output of the spatial-light modulator to an array of zone plates; and a moveable stage. Each pixel of the spatial-light modulator directs a beam of light to a specific zone plate, and each zone plate focuses that light to a spot on the substrate.

The spatial-light modulator must control the intensity of the light to a given zone plate in a quasi-continuous manner, from fully off to fully on. This 'gray-level scaling' enables one to control feature size (e.g. linewidths) to a much finer degree than the focal-spot diameter. The spatial-light modulator used in the initial research, and in the commercial ZPAL system, provides 256 intensity levels to 1089 beamlets.[17] As a result, feature-size control to the nanometer level appears feasible.[16] Gray-level scaling also enables compensation for the additive effects of overlapping tails of nearby



6.1 Schematic depiction of ZPAL, being commercialized by LumArray, Inc. A CW laser source illuminates a spatial-light modulator, each pixel of which controls the light intensity to one zone plate of the array. By adjusting the focal-spot intensity from zero to the maximum in a quasi-continuous manner, linewidth can be controlled and proximity effects corrected. By moving the stage under computer control, while intelligently modulating focal spot intensities, patterns of arbitrary geometry can be written.

focal spots, so-called proximity-effect correction, of which more is said in Section 6.5.

6.4.1 The use of diffractive-optical elements (DOEs)

The principle of operation of Fresnel zone plates is described in many texts on optics. In brief, an *amplitude zone plate* blocks light in alternate zones, leading to a first-order focal spot on-axis, as well as higher-order and virtual foci. Only 10% of the incident light arrives in the first-order focus. A *binary-phase Fresnel zone plate* (i.e., alternate zones introduce a pi-phase shift in the transmitted light) puts about 40% of the incident light into the first-order focus. By continuously varying the transmitted phase across each zone from 2π to 0 (i.e., a *blazed-Fresnel phase plate*) a focusing efficiency approaching 100% can be achieved. Zone plates exhibit chromatic and off-axis aberrations, but since ZPAL employs only narrow-band laser light and on-axis focusing, the focal spots should be free of these aberrations. Zone plates can operate at any wavelength; even neutral atoms have been focused with zone plates.[18] Zone plates can be made reproducibly using SEBL.[19] The NA, and hence the resolution (see Equation [6.1]) of a zone plate, is easily programmed, as is the focal length. The number of zones should be no larger than the reciprocal of the fractional bandwidth of the light source.

Zone plates are a subset of the much larger group of 'diffractive-optical elements'. By varying the zone progression, their shape, or the transmitted phase, one can do innovative wavefront engineering with diffractive-optical elements. For example, the light distribution in the focal spot can be altered so at to produce a central null at one wavelength and a central peak at another wavelength.[20] Such flexibility is not feasible with refractive lenses.

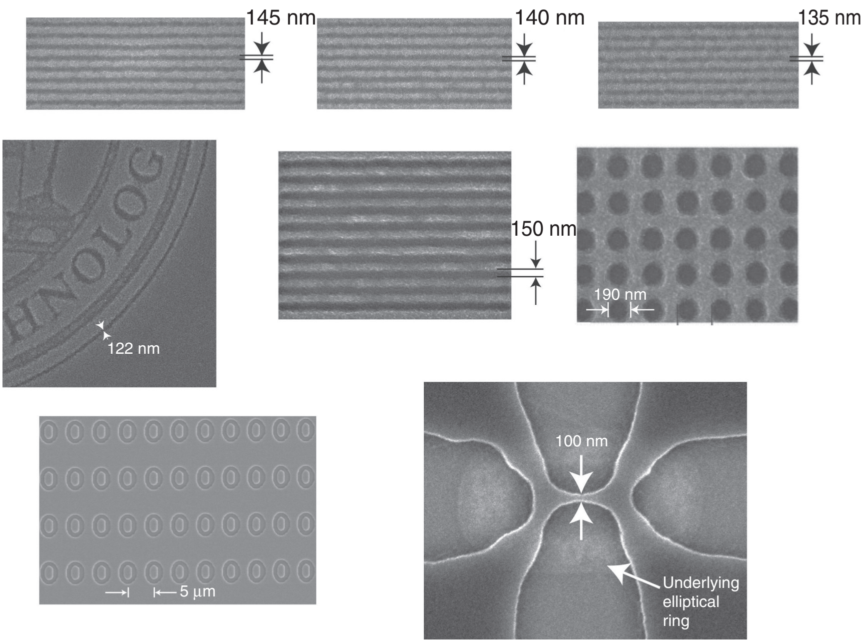
6.4.2 Scanning the stage as opposed to the beams

The strategy of scanning the stage has many advantages over scanning a beam. Stage position can be measured to within a fraction of a nanometer, using either a laser interferometer or an encoder. Beam scanning, whether with a beam of photons or charged particles, is more problematic. Encoder technology is highly robust and immune to disturbances, and can be made both precise and accurate. By scanning the stage to cover an entire substrate, the number of travel-direction reversals is minimized – an important consideration in maximizing throughput. The data that describe a desired pattern must be fractured and converted to appropriate form on a high-speed computer, and fed to the spatial-light modulator in coordination with the stage position.

The light distribution in the focal spots, the 'point-spread function' (PSF), of both Fresnel zone plates and refractive lenses, is approximated by an Airy function. Experimentally, it has been shown that for smooth line edges in resist, the overlap of focal spots should be at least 50%, i.e., an exposure-grid spacing of half the full-width at half-max of the PSF.[13–16] It is noteworthy that one can trade off the exposure-grid spacing, and hence the focal-spot overlap, with the number of gray levels from the spatial-light modulator. That is, the larger the overlap, the fewer are the number of gray levels required for a given linewidth control or proximity-effect correction.

6.4.3 ZPAL results in initial research at Massachusetts Institute of Technology (MIT)

Figure 6.2 illustrates the quality of the patterning achieved in early-stage research at MIT, using a 405 nm GaN laser source and binary-phase zone plates of 0.85 NA. Equation [6.1] indicates that the theoretical resolution



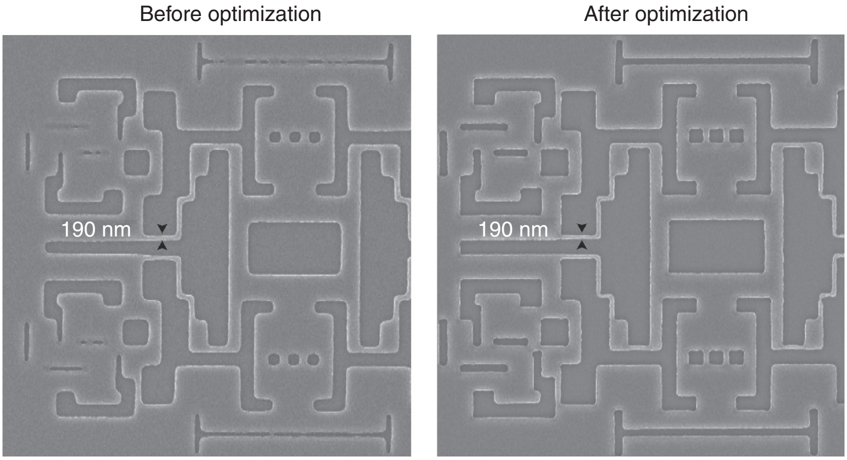
6.2 Scanning-electron micrographs of patterns written in photoresist using ZPAL, illustrating the facility with which patterns of arbitrary geometry can be written. The upper micrographs illustrate that a dense pattern of 135 nm half-pitch can be written using ZPAL and 405 nm radiation.

limit is p = 240 nm. Experimentally, a minimum period of 270 nm was achieved (half-pitch 135 nm). With water immersion, linewidths of 115 nm in dense arrays were achieved, using the 405 nm wavelength and binary-phase zone plates etched in Si_3N_4.[21]

6.5 Proximity-effect correction

Figure 6.3 illustrates the application of a proximity-effect-correction algorithm. Because *intensities* are added, not fields (i.e., linear superposition), the calculation and the simulation of patterns are greatly simplified, unlike the situation with partially-coherent optical-projection imaging.[9] Furthermore, the algorithm, which is based on an iterative nonlinear-optimization technique, converges to a solution in fewer than ten iterations. Being highly amenable to parallel processing, the resulting software yields results rapidly.[22]

As shown in Fig. 6.3, the corners of rectangles are sharpened and have a radius of curvature much smaller than the half-width-at-half-maximum



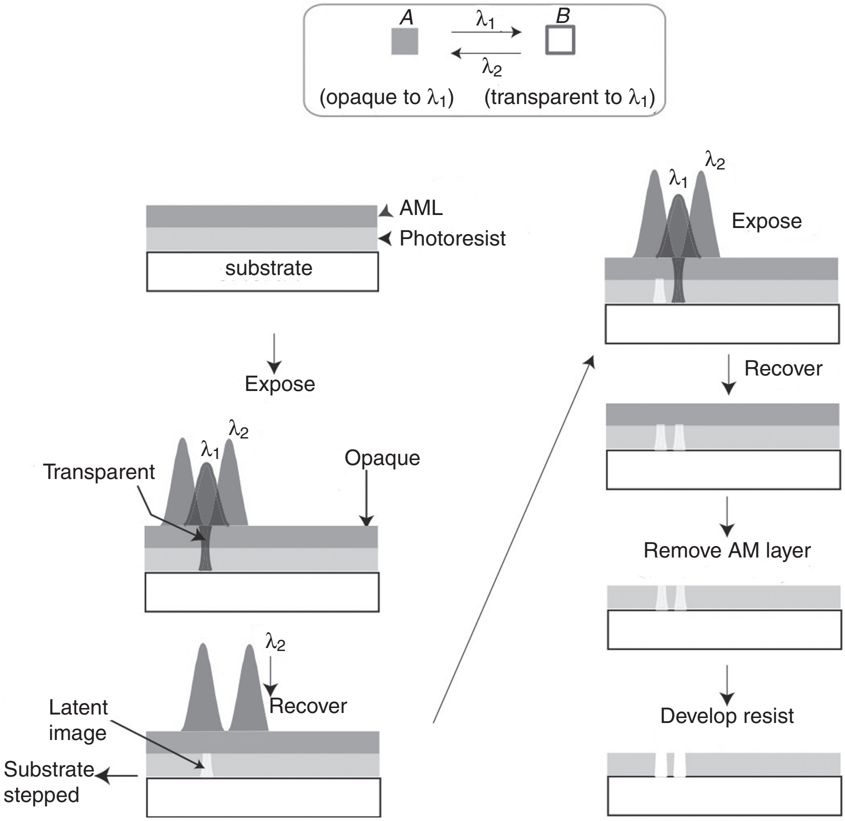
6.3 Optical micrographs of (left) a pattern as written without proximity-effect correction; (right) the same pattern written with correction. Note the sharpened corners of rectangles and the fidelity of exposure of small features.

(HWHM) of the focal spot's PSF. This is achieved by adjusting the intensities of the focal spots as the rectangle corners are approached. In the commercial ZPAL system, a general-purpose software code enables correction of optical-proximity effects over an entire pattern, irrespective of its complexity (http://www.lumarray.com).

6.6 Extending the resolution of ZPAL

Obviously, ZPAL could achieve even higher resolution than shown in Fig. 6.2 by employing a shorter wavelength. In fact, the first published description of ZPAL called for use of a wavelength of 4.5 nm, i.e., at the X-ray absorption edge of carbon, the primary constituent of most resists.[12] However, the limited availability and high cost of continuous sources of soft-X-ray radiation (e.g., undulators attached to a synchrotron) led to abandonment of this approach in favor of a novel invention based on photochromic chemistry – an approach called absorbance-modulation optical lithography (AMOL).[23–26]

Figure 6.4 depicts how an absorbance-modulation layer (AML), composed of appropriate photochromic molecules, can be used to squeeze the PSF of the light as it passes through the AML. The photochromic molecules would ideally have the property that they are converted from an opaque configuration to a transparent configuration by illumination of a short wavelength, λ_1, depicted as blue, and converted from the transparent configuration to the opaque one by illumination at a longer wavelength, λ_2, depicted


A
λ_1
B
λ_2
(opaque to λ_1)
(transparent to λ_1)
AML
Photoresist
substrate
Expose
λ_2
λ_1
Transparent
Opaque
λ_2
Recover
Latent image
Substrate stepped
λ_2
λ_1
Expose
Recover
Remove AM layer
Develop resist

6.4 Schematic of AMOL. The AML is composed of photochromic molecules. The short-wavelength light (λ_1) switches the molecules from their opaque configuration, A, to their transparent configuration, B, while the longer-wavelength light (λ_2) switches the molecules from B, back to A. By simultaneously illuminating the AML with a focal spot at λ_1 and a ring-shaped illumination at λ_2, the competition between the two photochemical reactions produces a sub-wavelength aperture through which the underlying photoresist is exposed by λ_1. The λ_1 light is then turned off enabling λ_2 to restore the AML to opacity. The stage is then moved to expose an adjacent feature, etc.

as red. Ideally, the AML should be thermally stable, fully reversible, and immune to quenching phenomena such as oxidation.

Simultaneous illumination of the AML with a focal spot at the short wavelength and a ring-shaped spot at the long wavelength, having a central null, results in the creation of a virtual aperture, through which the short-wavelength light, λ_1, can expose an underlying photoresist. Increasing the power at the long wavelength results in a further squeezing of the aperture. Simulations based on a rigorous model indicate that the virtual aperture can be squeezed below 20 nm. Preliminary experiments support the theoretical model.[24–26] Experiments by Ito et al., in which 20 nm features in periodic patterns were achieved using intimate-contact photolithography at the 365 nm wavelength,[27] lead one to expect that AMOL will be able to achieve resolution in the sub-20 nm region. We anticipate that tri-level resists and AR coatings will be necessary.

To accomplish the simultaneous illumination with λ_1 and λ_2, with the former as a normal focal spot and the latter as a ring-shaped spot with a central null, a unique diffractive-optical element, called a dichromat, was designed, fabricated, and tested.[20]

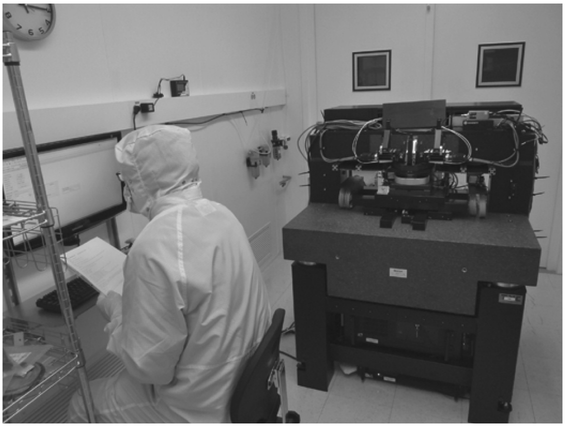
Further support for the feasibility of such extreme resolution with visible light comes from the research of S. Hell and colleagues, who achieved a resolution of 1/40 of the wavelength of light in fluorescence microscopy. A ring-shaped spot with a central null was used to deplete the excited state of fluorescent molecules, a process called stimulated emission depletion (STED).[8]

6.7 Commercialization of ZPAL by LumArray, Inc.

Following the initial research at MIT indicating the feasibility and utility of ZPAL, an effort to commercialize the technology was undertaken, resulting in the ZP-150A tool from LumArray, Inc. (see Fig. 6.5). It was designed to achieve a particular combination of throughput and resolution, emphasizing simplicity, flexibility, and affordability. The system specifications, shown in Table 6.1, make it suitable for researchers, and companies that manufacture custom products in small volumes.

Table 6.1 Specifications of LumArray's ZP-150

Minimum feature size	150 nm dense, 120 nm isolated
Numerical aperture	NA = 0.85
Parallel beams	1000
Writing speed	1.7 mm^2/s ~1 h per 100 mm wafer ~2 h per 150 mm wafer
Maximum pattern area	150 mm × 150 mm
Positioning resolution	1.2 nm
Design grid	1 nm
Overlay	20 nm
Field size	Unlimited
Exposure wavelength	405 nm
Mini-environment	ISO Class 5, 0.1°C
Pattern layout	GDSII
Tool size	89 cm × 135 cm × 155 cm



6.5 Photograph of LumArray's ZP-150 system. In order to show the system's internal parts and its compact size, the system's environmental enclosure has been removed.

6.7.1 Throughput

Throughput is expressed in Equation [6.2] above. In the ZP150A, $N = 1000$, $R = 290$ kHz, and $d = 75$ nm. Thus,

Areal writing speed, $A = 1.7$ mm^2/s.

If an entire 150 mm × 150 mm area is covered with a dense pattern requiring an 8-bit gray-level exposed pixel every 75 nm, the exposure data correspond to 4 terabytes. To achieve maximum efficiency in the data path, a customized field-programmable gate array (FPGA) is used to coordinate the data-delivery and synchronize with the scanning of the stage.

The zone plates in the ZP-150 are about 130 μm in diameter, arranged in six rows of 180 each. The rows are separated by about 3 mm so that the array covers an area of almost 24 mm × 20 mm, an area over which a typical wafer substrate can be held within the depth of focus by a pin chuck. For a wavelength of 405 nm and a NA of 0.85, the depth of focus is about half a micron.

6.7.2 Multilevel alignment

In order to align the writing relative to patterns already on the substrate, LumArray's ZP-150 uses an auxiliary set of zone plates operating in a confocal-imaging mode.[28] These zone plates are designed to have the same focal length as those used for writing, but they operate with red light so that photoresist is not exposed by the alignment operation.

6.7.3 Software

Software on the ZP-150 system consists of three major components: pre-exposure data preparation (PEDP), subsystem control and communication, and graphical-user interface (GUI). PEDP takes in the design layout, usually in GDSII format, performs a series of analyses on it, and generates an optimized data file that can be used by the ZP-150 for the exposure. The PEDP takes into account various system characteristics, such as the PSF of the focal spots, the associated proximity effects, and any misplacement of the individual zone-plate foci, and automatically applies corrections in the data. Because of the very large amount of data involved, the software is set up to take advantage of the computation power of a large computer cluster.

Architecturally, the ZP-150 system can be broken down to several subsystems: stage motion control, zone-plate-array-to-wafer gap control, and delivery of data between data storage and the spatial-light modulator. Each subsystem has its own set of software, and these must be coordinated and synchronized during system set-up and exposure. A set of communication protocols was developed, based on the standard Hypertext Transfer Protocol (HTTP). Since HTTP is very robust and operating-system independent, it meets overall system needs and minimizes development cost.

The GUI on ZP-150 is based on standard web interface and Java. Java provides a rich set of toolkits that facilitate rapid product development and provide a pleasant user experience. By adopting a web-based interface, the operator and the technician can remotely access the ZP-150 system. This is advantageous for system maintenance and repair.

6.8 Conclusion

In summary, the virtues of maskless ZPAL were first explored via academic research at MIT. Based on that, a commercial effort was undertaken, with the objective of providing the science and engineering communities with a low-cost, highly flexible, maskless photolithography system. Subsequently, the invention and demonstration of AMOL indicated that the resolution of ZPAL could be extended into the deep-sub-100 nm range without resorting to wavelengths shorter than the near-UV. Initial applications of the commercial ZPAL system will most likely be in research, development, and customized manufacturing. Improvements in ZPAL throughput can be achieved by increasing the number of pixels in the spatial-light modulator and its framing rate.

6.9 References

1. R. F. Pease (2005) 'Maskless lithography', *Microelectron. Eng.*, **78–79**, 381–392.
2. F. Pease (2010)'To charge or not to charge: 50 years of lithographic choices', *J. Vac. Sci. Technol. B*, **28**, C6A1–C6A6.
3. M. Mankos, H. F. Hess, D. L. Adler and K. J. Bertsche (2005) 'Maskless reflection electron beam projection lithography', US Patent 6,870,172 (KLA-Tencor Technol. Corp.)
4. Y. Zhang and P. Kruit (2007) 'High brightness 100-electron-beam source for high-resolution applications', *J. Vac. Sci. Technol. B*, **25**, 2239–2244.
5. J. T. Hastings, F. Zhang and H. I. Smith (2003) 'Nanometer-level stitching in raster-scanning e-beam lithography using spatial-phase locking', *J. Vac. Sci. Technol. B*, **21**(6), 2650–2656.
6. Y. Yang and J. T. Hastings (2007) 'Real-time spatial-phase locking for vector-scan electron beam lithography', *J. Vac. Sci. Technol. B*, **25**, 2072–2076.
7. H.I. Smith (1986) 'A statistical analysis of UV, X-ray and charged-particle lithographies', *J. Vac. Sci. Technol. B*, **4**, 148; H. I. Smith (1988) 'A model for comparing process latitude in ultraviolet, deep-ultraviolet and x-ray lithography', *J. Vac. Sci. Technol. B*, **6**, 346–349.
8. S. W. Hell (2007)'Far-field optical nanoscopy', *Science*, **316,** 1153.
9. D. Melville *et al*. (2010)'Demonstrating the benefits of source-mask optimization and enabling technologies through experiment and simulations', *Proc. SPIE*, **7640**, 764006.
10. A. H. Buxbaum, S. E. Fuller, W. Montgomery and M. E. Ungureit (2003)'Integration and optimization of the DUV ALTA pattern generation system using a CAR process with the Tetra photomask etch system', *Proc. SPIE*, **5256**, 1.
11. H. I. Smith (1999) Maskless lithography using multiplexed array of fresnel-zone-plates. US Patent #5900637, 4 May 1999; 'Microlens scanner for microlithography and wide-field confocal microscopy,' K. C. Johnson, US Patent #6,133,986, 17 October 2000.
12. H. I. Smith (1996) 'A proposal for maskless, zone-plate-array nanolithography', *J. Vac. Sci. Technol. B*, **14**, 4318–4322.
13. D. J. D. Carter, D. Gil, R. Menon, M. Mondol and H. I. Smith (1999) 'Maskless, parallel patterning with zone-plate array lithography (ZPAL)', *J. Vac. Sci. Technol. B*, **17**(6), 3449–3452.
14. Dario Gil, D. J. D. Carter, R. Menon, X. Tang and H. I. Smith (2002) 'Parallel maskless optical lithography for prototyping, low-volume production, and research', *J. Vac. Sci. Technol. B*, **20**(6), 2597–2601.
15. R. Menon, A. Patel, E. E. Moon and H. I. Smith (2004) 'Alpha-prototype system for zone-plate-array lithography', *J. Vac. Sci. Technol. B*, **22**(6), 3032–3037, Nov/Dec (2004).
16. H. I. Smith, R. Menon, A. Patel, D. Chao, M. Walsh and G. Barbastathis (2006) 'Zone-plate-array lithography: a low-cost complement or competitor to scanning-electron-beam lithography', *Microelectron. Eng.*, **83**, 956–961.
17. A. Payne, W. DeGroot, R. Monteverde and D. Amm (2004) 'Janis spatial light modulator from silicon light machines', *Proc. SPIE*, **5348**, 76.

18. T. Reisinger, Sabrina Eder, M. M. Greve, H. I. Smith and B. Holst (2010) 'Free-standing silicon-nitride zoneplates for neutral-helium microscopy', *Microelectron. Eng*., **87**, 1011–1014.
19. Dario Gil, R. Menon and H. I. Smith (2003) 'Fabrication of high-numerical-aperture phase zone plates with a single lithography exposure and no etching', *J. Vac. Sci. Technol. B*, **21**(6), 2956–2960.
20. R. Menon, P. Rogge and H. -Y. Tsai (2009) 'Design of diffractive lenses that generate optical nulls without phase singularities', *J. Opt. Soc. Am. A*, **26**(2), 297.
21. D. Chao, A. Patel, T. Barwicz, H.I. Smith and R. Menon (2005) 'Immersion zone-plate-array lithography', *J. Vac. Sci. Technol. B*, **23**, 2657–2661.
22. R. Menon (2003) PhD Thesis, 'Diffractive Optics for Maskless Lithography and Imaging', Massachusetts Institute of Technology, June 2003.
23. R. Menon and H. I. Smith (2006) 'Absorbance-modulation optical lithography', *J. Opt. Soc. Amer. A*, **23**, 2290–2294.
24. R. Menon, H. -Y. Tsai, W Samuel and Thomas III (2007) 'Far-field generation of localized light fields using absorbance modulation', *Phys. Rev. Lett*., **98**, 043905.
25. T. L. Andrew, H. -Y. Tsai and R. Menon (2009) 'Confining light to deep sub-wavelength dimensions to enable optical nanopatterning', *Science,* **324**, 917.
26. N. Brimhall, T. L. Andrew, R. V. Manthena and R. Menon (2011) 'Breaking the far-field diffraction limit in optical nanopatterning via repeated photochemical and electrochemical transitions in photochromic molecules', *Phys. Rev. Lett.*, **107**, 205501.
27. T. Ito, T. Yamada, Y. Inao, T. Yamaguchi, N. Mizutani and R. Kuroda (2006) 'Fabrication of half-pitch 32 nm resist patterns using near-field lithography with a-Si mask', *Appl. Phys. Lett*., **89**, 033113.
28. R. Menon, E. E. Moon, M. K. Mondol, F. J. Castano and H. I. Smith (2004) 'Scanning-spatial-phase alignment for zone-plate-array lithography', *J. Vac. Sci. Technol. B*, **22**(6), 3382–3385, Nov/Dec (2004).

7

Chemistry and processing of resists for nanolithography

A. NOVEMBRE, Princeton University, USA and
S. LIU, IBM Corporation, USA

DOI: 10.1533/9780857098757.194

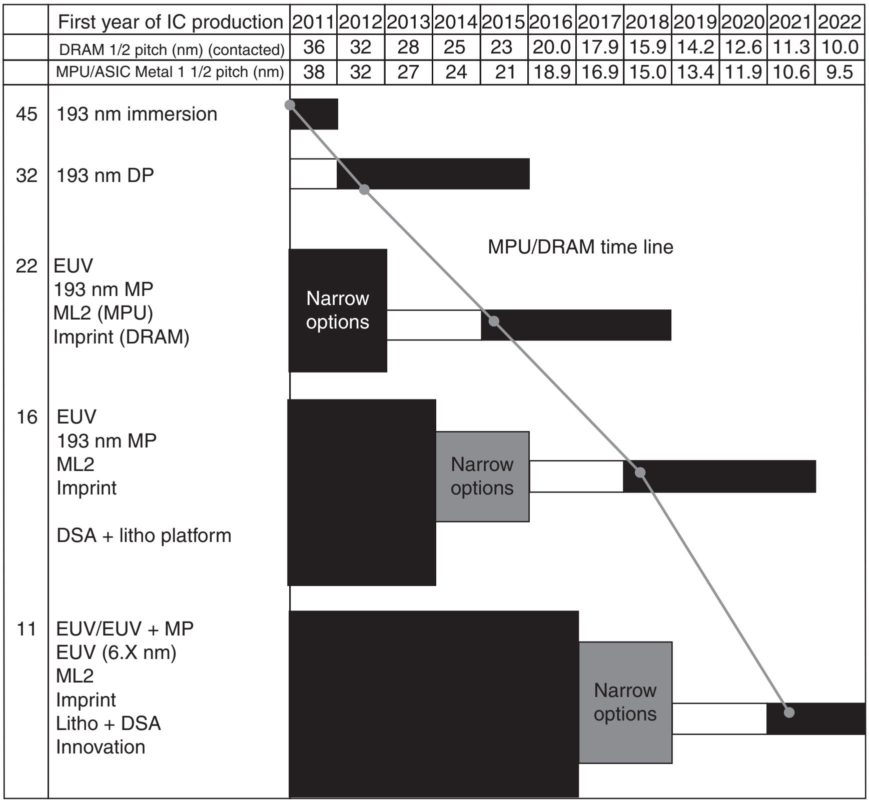
Abstract: This chapter serves as a review of the synthesis and radiation induced chemistry of resist materials and applicable steps used in pattern delineation processes. Fundamental requirements and limitations influencing the choice of a resist for a particular lithography exposure technology are described. An extensive discussion of the use of chemically amplified resist platforms in the current and proposed states of art lithography processes is also highlighted. Suitable resist materials for non-radiation-based lithography processes are also presented as part of the discussion on forward looking lithography technologies.

Key words: chemically amplified resists, inorganic resists, photo acid generators, radiation chemical yields, resists for double patterning and double exposure processes, EUV resists, e-beam resists, immersion lithography compatible resists.

7.1 Introduction

The intense drive towards designing and fabricating integrated circuits having a minimum element dimension of < 50 nm and an increasing density of the elements in a semiconductor material represents a truly global effort.[1] This trend for the dynamic random access memory device (DRAM) and microprocessor (MPU) is represented in Fig. 7.1. The ability to shrink the critical feature size of the device is dependent upon the lithographic technologies involved in the delineation of the circuit pattern.[2] The fundamental lithographic process using a photon-based exposure source is shown in Fig. 7.2. This process consists of the delineation of the desired circuit pattern in a material referred to as resist, followed by transfer of the pattern using an appropriate additive (deposition) or subtractive (etching) technique.[3]

Historically, resist materials are amorphous organic and organometallic polymers readily soluble in both aqueous and organic solvents and containing small molecule additives that serve to enhance the lithographic



7.1 Minimum DRAM and MPU device feature size vs the first year of integrated circuit (IC) production. (*Source*: Semiconductor Industry Association. The International Technology Roadmap for Semiconductors, 2011 Update. SEMATECH: Austin, TX).

performance of the material.[4] A subset of resists is derived from molecular and inorganic based materials.[5–7] The primary requirements of the resist as part of the multi-step lithographic process are: (i) the material must have a high susceptibility to undergo a radiation (exposure) induced chemical transformation; and (ii) must maintain its structural integrity in the environment used during the subsequent lithographic pattern transfer step.[8]

The following list summarizes the primary (radiation) and secondary (non-radiation) lithographic patterning processes:

Primary type: radiation based:

- *Exposure source*: photons – 436 nm (Hg g-line), 365 nm (Hg i-line), 248 nm (KrF Laser), 193 nm (ArF Laser), 157 nm (F_2 Laser), 13.5 nm (EUV), 0.5–1.5 nm (X-Ray)

- *Exposure source*: charged particles – electrons, ions (Ga, He, H) 1–100 keV

Secondary type: non-radiation based:

- Nanoimprint
- Dip-pen/polymer pen
- Directed self-assembly.

In conjunction with each of the above lithographic technologies, resists act as a key enabler in increasing the density of the discrete electronic elements in a given area of an integrated circuit (IC). However, the effective size (length scale) of some of the critical components of the resist formulations does pose limits on their enabling capability. With these boundary conditions defined, this chapter serves as a review of the synthesis, radiation induced chemistry, and applicable process steps enabling the use of a resist material as part of the IC pattern formation process.

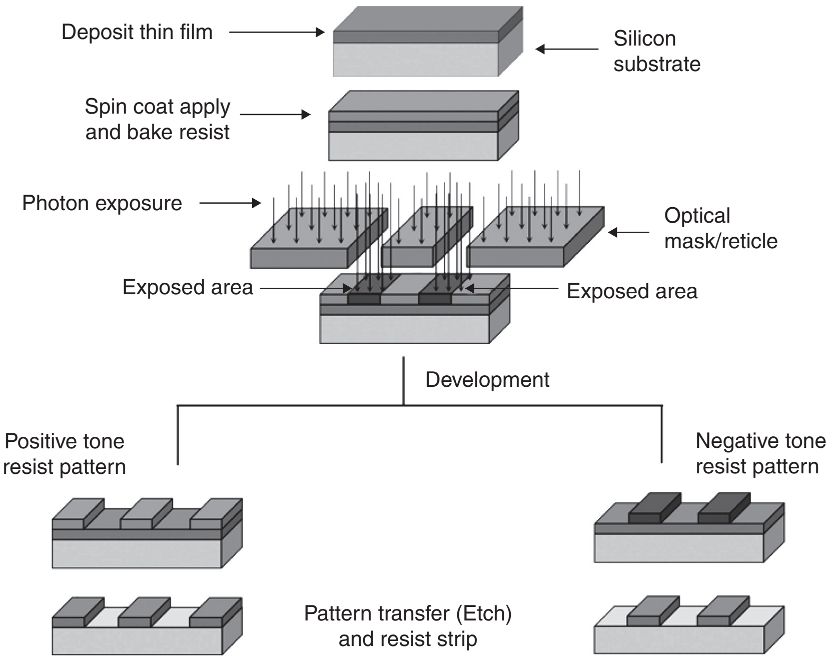
7.1.1 Resist design considerations

As depicted in Fig. 7.2, the pattern or relief image in the thin film of a resist is principally fabricated through the action of the incident radiation of a lithography tool. The exposing radiation initiates chemical events which produce a change in the solubility in a developing agent for the exposed vs non-exposed regions of the resist film. This differential solubility is used to produce two types of relief images, whereby resists are classified as being either positive or negative acting. A positive acting resist is one in which the exposed areas are removed during development while, for a negative acting resist, the exposed areas remain after development.

The exposure source and strategy in commercial lithography tools and the type of resist used as a function of the minimum resist feature size are provided in Table 7.1.

Concurrent with the use of photons as the exposure source are charge particles such as electrons and ions, which have been exploited to achieve one of the highest levels of resolution observed by a resist system.[9] Charge particle-electron based lithography exposure systems are principally involved in the fabrication of masks (reticles) used in photon-based lithography tools, and are more extensively described in Chapter 3.

In practical terms, minimizing the time required to expose the resist pays back in terms of improving throughput and production costs. Therefore, an essential requirement in the design of a resist is that the radiation


Deposit thin film
Silicon substrate
Spin coat apply and bake resist
Photon exposure
Optical mask/reticle
Exposed area
Exposed area
Development
Positive tone resist pattern
Negative tone resist pattern
Pattern transfer (Etch) and resist strip

7.2 Process flow diagram for the resist pattern delineation process.

Table 7.1 List of the exposure technology and type of resist used to delineate a minimum resist feature size ranging from <0.10 to >1.0 μm

Min. feature Size (μm)	Exposure source & strategy Tool/λ(nm)	Resist type
1.25 – 0.60	g-line/436	Novolac/DNQ
0.60 – 0.40	i-line/365	Novolac/DNQ
0.40 – 0.25	i-line\PSM DUV/248	Novolac/DNQ, Chemically Amplified (CA), Organometallic (Si)
0.25 – 0.10	DUV/248\PSM, DUV/193\immersion(i) e-beam, X-ray/0.5-1.5	CA, Block Copolymers Organometallic (Si), inorganic
≤0.10	DUV/193i\multiple patterning X-Ray/0.5-1.5,e-beam EUV/13.5, Non radiation based	CA, Block Copolymers, Inorganic, Molecular – small molecule

induced chemical transformation must be highly efficient, ensuring that the incident exposure dose threshold is a minimum. This in turns defines the sensitivity of the resist to the form of radiation being used in the lithography exposure system.

Table 7.2 Required resist sensitivity/process dose vs lithography exposure technology

Lithography exposure technology	Resist sensitivity/Process dose
248 nm	10–50 mJ/cm^2
193 nm	10–50 mJ/cm^2
Extreme ultraviolet (13.5 nm)	5–20 mJ/cm^2
X-ray (0.5–1.5 nm)	5–20 mJ/cm^2
High keV (50–100) e-beam	5–30 μC/cm^2
Low keV (1–5) e-beam	0.2–1.0 μC/cm^2

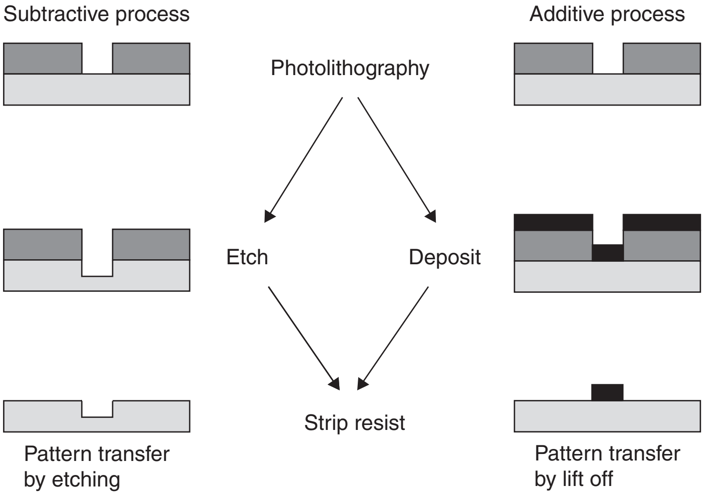
7.1.2 Influence of the lithography exposure system on required resist sensitivity and performance

As discussed in the previous section, various lithographic technologies have been developed to meet the ever continuing demand of increasing the pattern density of an IC. Table 7.1 summarizes the options for the type of resists that have been investigated for a particular exposure tool. Correspondingly, Table 7.2 represents the resist sensitivity/process dose required for high throughput patterning as a function of the lithography exposure technology.

A further consideration when designing a resist for a particular lithography exposure system is the radiation induced chemistry must be compatible with the operating conditions and writing strategy of the exposure tool.[10] For any lithography exposure system, outgassing of components of the resist film while present in the exposure tool can be detrimental to its operation, performance, maintenance, and useful lifetime.[11]

7.1.3 Stability of the resist in the pattern transfer environment

Figure 7.3 details processes used to transfer the resist pattern into the underlying substrate layer. In the subtractive process the resist acts as a mask for areas of the underlying layer, whereas in the additive process it serves as a stencil to enable the deposition of a material in the areas where the underlying layer is exposed to the deposition environment.[12] The key resist requirements in either process are to maintain structural and mechanical integrity when subjected to the chemicals used to produce the pattern in the underlying layer, and removable after completion of the pattern transfer step. In the subtractive plasma (physico-chemical) dry etching environment, the resist must withstand particle bombardment and plasma-induced chemical reactions during the time the plasma is etching away the underlying layer.[13] The standard measure of etching resistance is the selectivity defined as the etch rate of the material being etched vs the etch rate of the masking material or substrate.[14]



7.3 Process flow for a subtractive and additive resist pattern transfer step.

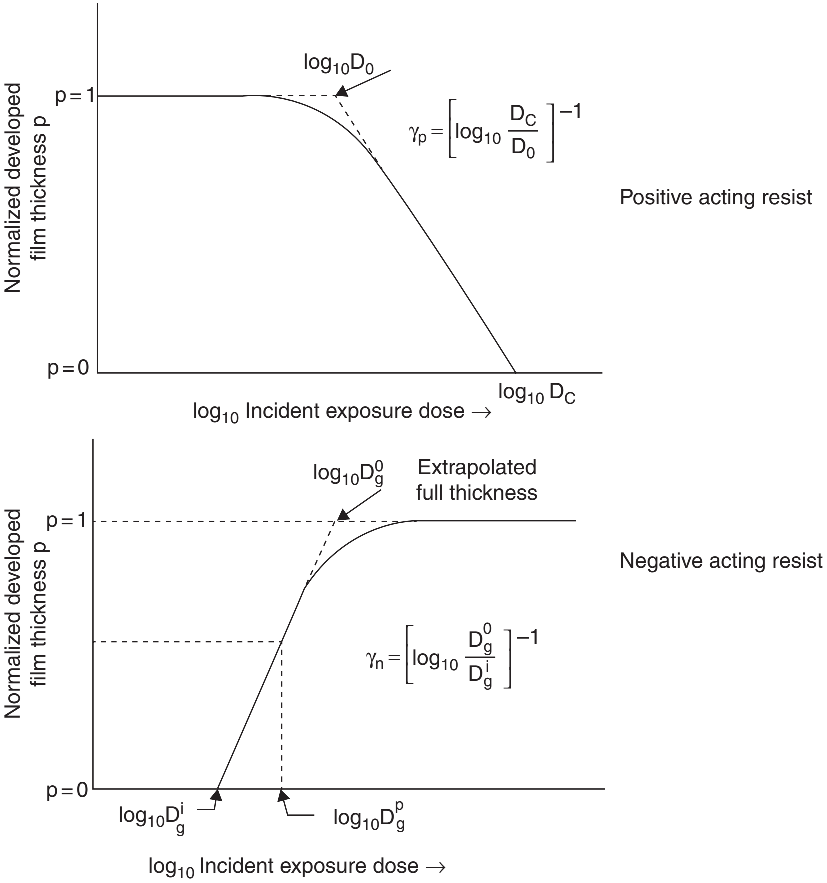
Higher selectivity values afford the opportunity to make use of thinner initial resist film thicknesses, which in turn can promote the resolution capability of the resist material.

In the additive process a metal, typically Aluminum or Gold, is deposited on top of both the patterned resist layer and in between the resist relief images. In the subsequent step the resist is then removed or lifted off from the base substrate, leaving only metal patterns initially deposited in areas where the base substrate was exposed to the deposition environment.[15]

7.1.4 Resist performance characterization methodologies

The quantitative analysis and means of comparison of the performance of resist materials is initially obtained through generation of an exposure response curve. The curve provides for the determination of the type of resist (positive or negative), the intrinsic exposure sensitivity and contrast of the material. Exposure response curves are generated by plotting the normalized resist film thickness remaining after development (the remaining thickness is normalized to the initially coated resist thickness) vs the log of the incident exposure dose. Figure 7.4 provides generic exposure (contrast) curves for a positive and a negative acting resist.

The contrast (Υ) is calculated from the slope of the line tangent to the linear segment of the resist exposure response curve.[16] Contrast is critically

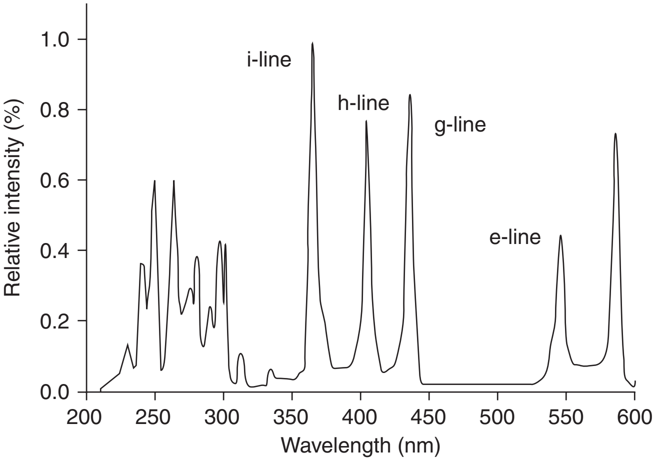


7.4 Generic exposure response/contrast (γ) curves for positive and negative acting resists.

dependent on the solubility discrimination mechanism imparted in the resist film as a consequence of the radiation induced chemistry. In general, higher contrast results in higher inherent resolution capability of the resist and numerous points along the contrast curve have been used as a measure of the resist sensitivity. The various resist contrast mechanisms will be described in the following sections of this chapter.

7.2 Resists for optical lithography: synthesis and radiation induced chemistry of resists as a function of exposure technology

Modern day lithographic resists can trace their origins back to the printing industry. Early version resists were of the negative acting type, resulting



7.5 Spectral output of a typical high pressure Hg arc lamp used in commercial optical (i and g-line) lithography exposure tools.

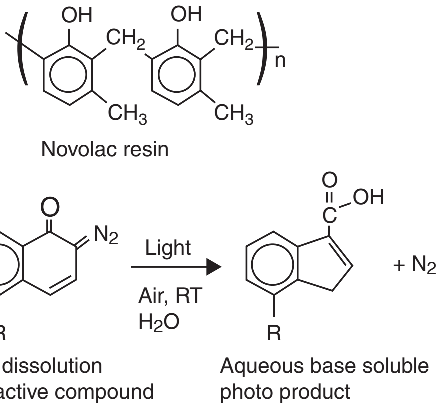
from a photoinduced crosslinking reaction within the material leading to an insoluble network. An informative perspective of early photoresists is provided by Reiser.[17] Specialized and tailored resist formulations evolved with the commercialization of photolithography tools employing a high pressure Mercury arc lamp as the exposure source.[18] Figure 7.5 provides the emission spectra for the Hg arc lamp.

7.2.1 Positive acting photo resists for g- and i-line lithography

The introduction of 436 and 365 nm (g- and i-line) optical step and repeat reduction steppers led to the custom development and use of positive acting photoresists consisting of a diazonaphthoquinone photoactive compound (DNQ-PAC) and a novolac resin.[3] The chemical structure representation for the components of the photoresist and the photo-induced chemistry involved in the formation of positive tone images is provided in Fig. 7.6.

The contrast mechanism for this family of resists is based on the photo exposure inducing a change in the chemical polarity/hydrophobicity of the exposed vs the non-exposed regions of the film. The primary function of the DNQ-PAC is to inhibit the dissolution of the novolac resins in an aqueous alkaline developer solution. The structure of the DNQ-PAC is tailored to ensure optimal film absorption to produce a vertical resist pattern profile.

The predominant synthetic pathway used in the preparation of the DNQ-PAC is through the esterification reaction of a phenol based ballast compound with a DNQ sulfonyl chloride.[19] From a thin film formation


OH
CH2
OH
CH2
n
CH3
CH3
Novolac resin
O
N2
Light
Air, RT
H2O
R
O
C
OH
+ N2
R
Aqueous base dissolution inhibitor/photoactive compound
Aqueous base soluble photo product

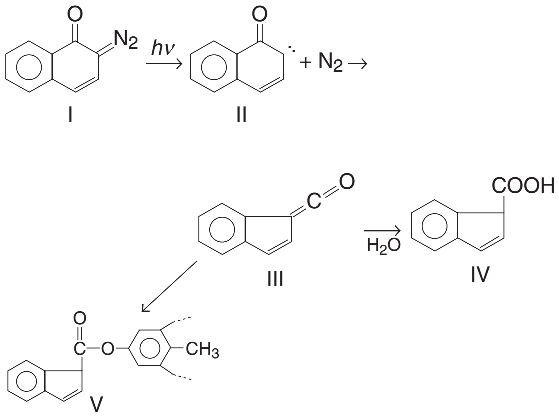
7.6 Chemical structure representation of the components in a standard photoresist formulation and the reaction sequence for the photoinduced chemistry occurring in these materials.

perspective, it is critical the DNQ-PAC is compatible with the novolac resin, thereby ensuring a single phase within the film. The phase compatibility is essential for providing dissolution and uniform development of only the exposed area of the film.

The dissolution inhibition efficiency of the DNQ-PAC when present in a novolac resin has been shown to be dependent on the number of DNQ moieties per photoactive compound, on the degree of esterification for polyfunctional DNQ-PACs, and the location of each moiety within the DNQ-PAC molecule.[20] In addition, when investigating the interaction of the DNQ-PAC with the novolac, the degree to which it can form hydrogen bonds with the resin and the degree of its hydrophobicity also play important roles in the solubility discrimination observed after the film is exposed and developed.[21]

In parallel with the optimization of the DNQ-PAC was the search to identify the ideal structure and properties of the novolac resin. The novolac resins are typically synthesized via a condensation reaction of formaldehyde and cresol isomers to produce a relatively low molecular weight polymer. The weakly acidic phenolic structure of the resin in the absence of the DNC-PAC is soluble in an aqueous alkaline solution. The molecular weight, molecular weight distribution, molecular (isomeric) structure of the starting cresol monomer and the position of the methylene bond within the polymeric chain influence the performance of these materials as resists for 365–436 nm wavelength photolithography.[22] Figure 7.7 provides for the predominant photo-induced reaction pathways of the DNQ-PAC in the resist.

Exposure to light initially leads to elimination of nitrogen to produce an intermediary carbene (II). The carbene undergoes a Wolff rearrangement to form a ketene (III). In the presence of small quantities of water, the ketene



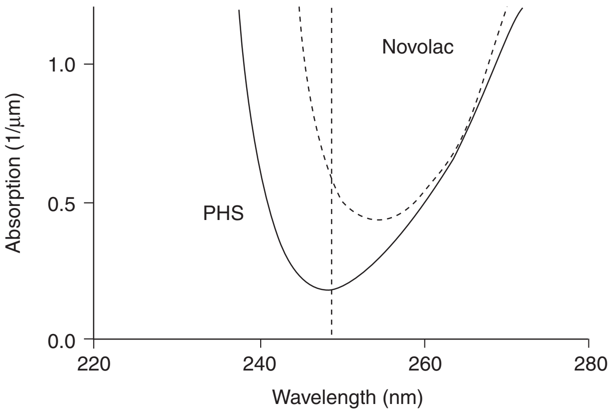
7.7 DNQ-PAC photolysis reaction sequence.

is converted to an aqueous base soluble indene carboxylic acid (IV). In the absence of moisture, the ketene reacts with the resin to form a phenol ester (V), which serves to induce the crosslinking in the film. This alternative path (V) is highly undesirable when attempting to form a positive tone pattern in the photoresist.[23,24]

An additional desirable property of the DNQ-PAC is a reduction in its absorption at the exposure wavelength as the exposure process proceeds in the film.[25] The bleaching of the film ensures uniform exposure throughout the film, resulting in near vertical positive tone image pattern profiles. DNQ-novolac-based photoresists also exhibit minimal developer induced erosion of the non-exposed regions of the film, along with no evidence of pattern deformation resulting from swelling of the image during the development process. Finally, these resists afford the ability to use a non-flammable water-based developing solvent, which eases the handling and removal of the spent developer and promotes a safer manufacturing environment.

7.2.2 Chemically amplified resists for 248 nm lithography

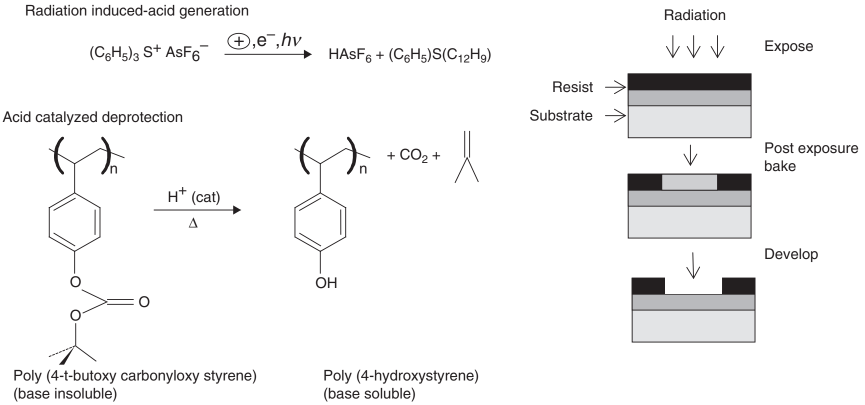
The practical limit of conventional (i-line) photolithography to print resist patterns below 300 nm led to the pursuit of building an optical lithography exposure tool operating at a wavelength below that of the Hg arc lamp i-line emission line. This new generation of optical lithography tools employed a photon source based on an excimer laser. The first generation Krypton Fluoride (KrF) laser emitted light at $\lambda = 248$ nm and, through the requirements of incorporating line narrowing elements into the illumination system, resulted in an appreciable reduction in the light available to expose the resist.[26]



7.8 UV-Visible absorption spectra of novolac and polyhydroxystyrene polymers. The dotted vertical line is shown to provide the absorption of the two polymers at λ = 248 nm.

Investigation of the use of the conventional i-line lithography DNQ-novolac resists with the KrF laser lithography tool revealed the immediate shortcomings of these materials. The conventional photoresists were found to be too highly absorbing/opaque in the 250 nm wavelength regime and did not bleach as the exposure progressed.[27] Figure 7.8 shows the 220–280 nm absorption spectrum for a conventional novolac resist. Novolac-based resists are also characterized by having relatively low intrinsic sensitivity and quantum yield values for the photolysis of the DNQ-PAC. These combined effects limited the ability of these resists to be used to pattern vertical images in a 1.0 μm thick film at a low exposure dose. These adverse resist properties necessitated a paradigm shift in the development of resists and the corresponding radiation induced chemistry that would be used in conjunction with 248 nm and below generation lithography technologies.

The introduction of resists based on chemical amplification represented the most viable approach to meeting the high sensitivity/throughput demand of DUV lithography exposure tools.[28] The predominant chemical amplification mechanism employed in DUV resists is based on a single photo-induced absorption event initiating a multiplicity of subsequent chemical reactions.[29] Numerous research groups utilized this approach by incorporating an acid generating species in a polymer matrix containing an acid labile group.[30–32] Figure 7.9 is a schematic representation of the radiation induced chemical amplification mechanism for a two-component positive acting resist, along with the corresponding processing sequence used in the patterning of the resist. In this example, which will be more extensively discussed in Section 7.2.3 of this chapter, the acid generating species is a triphenylsulfonium hexafluoroarsenate onium salt, and the polymer is


Radiation induced-acid generation
$(C_6H_5)_3\,S^+\,AsF_6^- \xrightarrow{\oplus, e^-, h\nu} HAsF_6 + (C_6H_5)S(C_{12}H_9)$
Acid catalyzed deprotection
n
H^+ (cat)
Δ
O
O
O
n
+ CO_2 +
OH
Poly (4-t-butoxy carbonyloxy styrene)
(base insoluble)
Poly (4-hydroxystyrene)
(base soluble)
Radiation
Expose
Resist
Substrate
Post exposure bake
Develop

7.9 Schematic representation of the two-step chemical amplification process and corresponding process sequence when using an onium salt as the acid generating species and a 4-t-butoxy-carbonyloxy protected polyhydroxystyrene polymer.

poly(4-tert-butoxycarbonyloxystyrene) (PTBOC).[28] The tert-butoxycarbonyloxy (TBOC) group is used to affect differential solubility between the exposed and non-exposed regions of the film or, in other terms, to protect the non-exposed film regions of the polyhydroxystyrene (PHS)-based polymer from dissolving in an aqueous base developer solution. As shown in Fig. 7.8, PHS based polymers exhibit a higher degree of optical transparency vs a novolac resin, thereby afford a more uniform exposure through the desired depth of the resist film and vertical pattern profiles. The remaining component, the onium salt, is the photoactive compound in which a small amount, upon exposure to DUV light, is converted to the strong Brønsted acid hexafluoroarsenic acid.[33] In a subsequent post-exposure heating (film baking) step, the acid catalyzes the thermolysis of the TBOC group present on the polymer. This leads to the conversion of the hydrophobic PTBOC polymer to its polar PHS form, the liberation of gaseous products and the regeneration of the initial acid molecule. Amplification is said to occur whereby a single acid molecule can be responsible for removal of 100 or more TBOC groups. This quantitative measure of the removal of protecting groups per acid molecule is also referred to as the catalytic chain length for the chemically amplified (CA) resist system.[29] The high degree of chemical amplification, in combination with the proper post-exposure bake conditions, enables the use of a very low incident exposure dose (10–50 mJ/cm^2) as part of the imaging process.[34] The ability to use the desirable industry standard 0.26 N TetraMethylAmmonium Hydroxide (TMAH) aqueous alkaline developer common with DNQ-novolac resists is also preserved with this approach. Conversely, if an organic based developer is used, negative tone patterns can be formed leading to a bi-tonal CA resist material.[35] The following sections detail the types of acid generating species, the polymer matrix materials and the pendent acid labile groups used in a resist to maximize the efficiency of the radiation induced chemical amplification process.

Acid generator chemistry

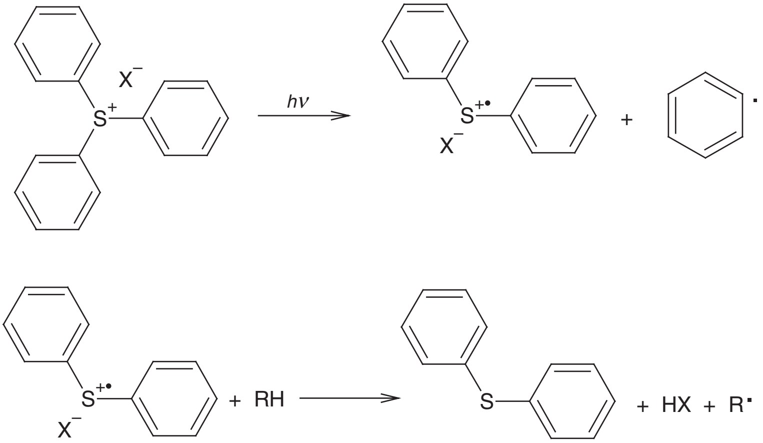
The initial latent image step in resists based on chemical amplification is the radiation induced generation of an acid from its precursor. The acid precursor can exist as a small molecule additive, as a moiety or side chain group on the matrix polymer, or part of the backbone structure of the polymer.[36–37] The exposure generated acid has been used to formulate positive acting resists based on a chemical polarity change of the non-exposed vs exposed film regions, and negative working resists based on insolubilization of the exposed areas resulting from crosslinking of the backbone polymer network of the matrix polymer.[38]

Radiation generated acids are of two primary types: ionic and nonionic. The key requirements that must be considered when designing and selecting the appropriate acid generator species are:

- The absorption properties must be tuned to the exposure wavelength of the lithography tool.
- There must be structural flexibility to produce a broad range in the strength of the radiation generated acid.
- The quantum efficiency of acid formation should be high such that the initial quantity of the acid precursor present in the film is low. This serves to reduce the acid precursor acting as an aqueous base dissolution inhibitor and potentially driving up the required process exposure dose.
- The acid precursor when present as a separate species should be compatible with the other components of the CA resist formulation.
- The acid precursor and the radiation generated acidic product must not become volatile during the exposure process and in the temperature regime used in the process baking steps.
- The radiation generated acid should be constructed so that diffusion of the species outside of the exposed areas of the film is a minimum.
- The manufacturability and cost of the acid precursor material need to be considered.

Ionic acid generating species

The ionic acid generating species, gaining wide attention at the time the IC industry began to propose the need to move to using DUV lithography exposure tools, was based on a class of materials referred to as onium salts



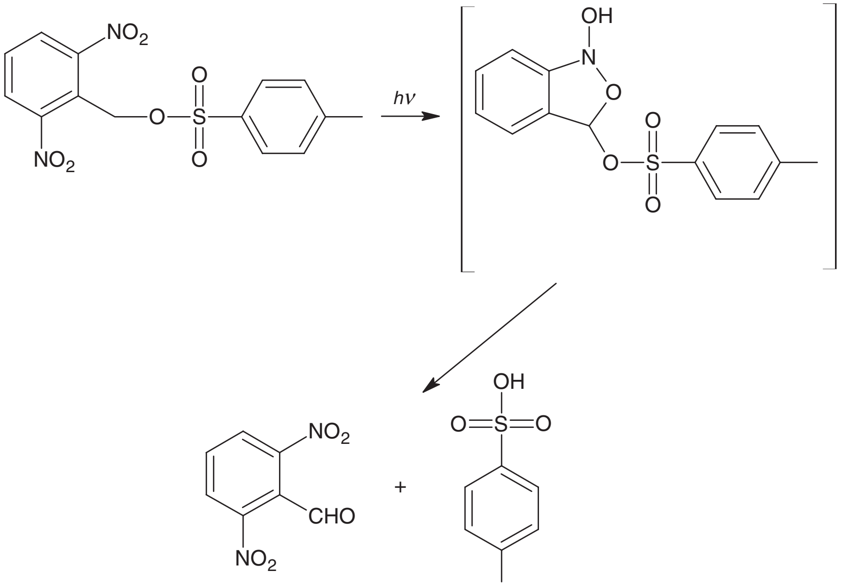
7.10 Generalized DUV photolysis mechanism for triphenylsulfonium based ionic acid generating species.

developed by Crivello.[33] Iodonium and Sulfonium salts were extensively investigated, and were found to exhibit high measured quantum yields and flexibility in varying the strength of the radiation generated acid.[39] In addition, proper choice of the counter ion minimized the volatility of the acid and diffusion into the non-exposed regions of the film. For example, the use of the methanesulfonate anion represented an exciting choice in meeting the properties just described.[40] As will be discussed in the section on 193 nm lithography resists, spectral response tuning, through choice of the aromatic or aliphatic substituents present on the ionic acid precursor, extended the use of these materials in 193 nm and below lithography technologies.[41,42]

The postulated generalized photochemistry of sulfonium salts is provided in Fig. 7.10. Photolysis initially generates a reactive radical, radical cation, and cation intermediates. The second step of the process involves hydrogen abstraction from the surrounding medium to produce the Brønsted acid.[43] The formation of the radical and cationic species have also found utility in curing reactions suitable for use in negative acting resists.[44]

Nonionic acid generating species

Radiation induced acid generation can also be achieved starting with nonionic acid precursor materials, such as those based on 2-nitrobenzyl esters.[45,46] These precursors undergo intramolecular o-nitrobenzyl rearrangement upon exposure, leading to the formation of toluenesulfonic acid. The photolysis mechanism for the 2-nitrobenzyl ester acid precursor is shown in Fig. 7.11. This acid precursor lacks the ability to generate radical species during exposure and effectively shuts down the reaction pathway, leading to



7.11 DUV photolysis and acid formation mechanism for 2-nitrobenzylesters.

unwanted competing crosslinking reactions.[47] A host of acid precursor materials including o- and p-nitrobenzyl sulfonic acid esters,[48] alkylsulfonates,[49] alpha sulfonyloxyketones,[50] iminosulfonates,[51] and aryl methyl sulfones[52] have been shown to generate photo-induced strong to moderate acids.

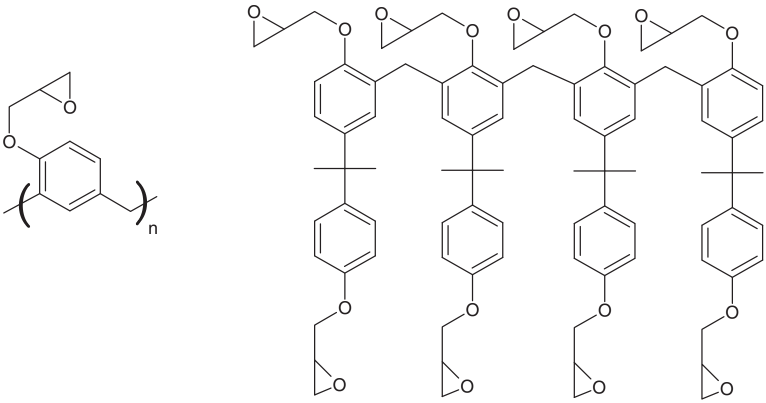
Nonionic acids also have an added desirable characteristic of being compatible with a broad range of matrix polymer resins and, through structural modifications, are able to be used in 193 nm and beyond lithography resists.

7.2.3 Acid catalyzed reactions

The acid generated during the exposure step has been shown to be highly versatile in catalyzing reactions, which has enabled the development of both positive and negative acting resists. The following sections describe a number of reaction pathways used to delineate a particular tone of resist.

Negative acting chemically amplified resists

The predominant radiation induced reaction pathways utilized to create negative acting CA resists include cationic and condensation polymerizations, electrophilic aromatic substitution, and acid catalyzed rearrangement.[4,20] The negative resist formulations are typically multi-component and make use of either an ionic or nonionic acid generating species.

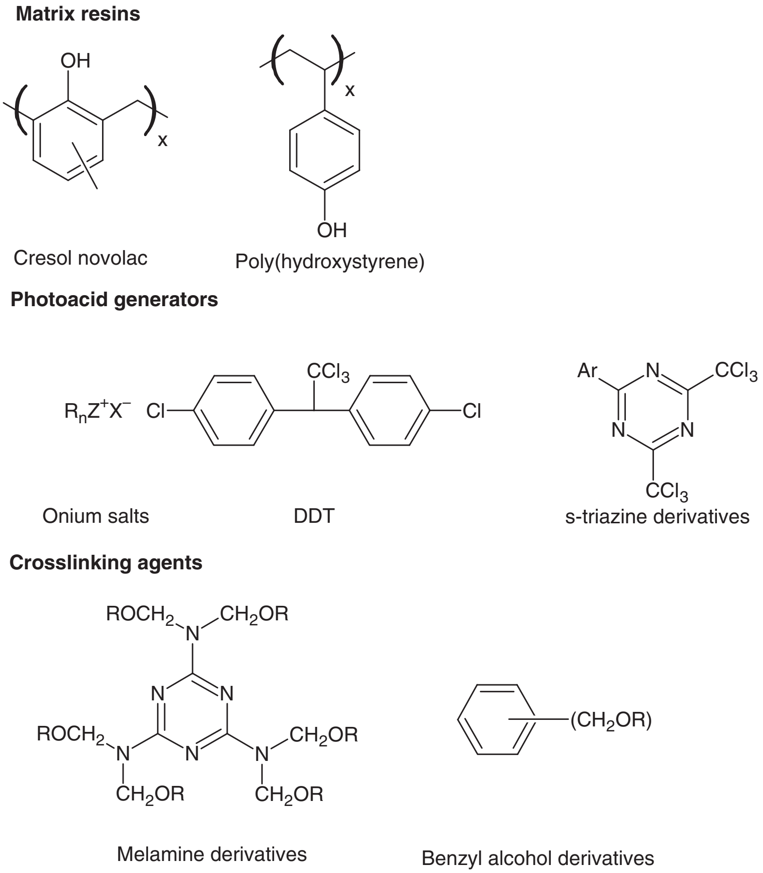


7.12 Chemical structure of epoxy-containing resins used in negative acting CA resist formulations.

Cationic polymerization reaction pathway

Use of the exposure induced acid catalyzed cationic polymerization reaction pathway represents one of the earliest examples of a negative resist system based on chemical amplification. Specifically, this involved the acid catalyzed ring opening polymerization of epoxy-containing resists.[53,54] Representative structures of epoxy-containing resins used in these resist formulations are provided in Fig. 7.12.

The exposure generated acid initiates the opening of a polymer bound epoxide ring, which induces a reaction with an epoxide ring of another polymer chain. This reaction results in the formation of an inter-chain crosslink, which continues between subsequent chains and leads to an insoluble polymer network in the exposed film regions.[55] These resist systems were highly sensitive to most forms of radiation used in lithography exposure tools. They typically required a low exposure dose to approach complete insolubilization of the exposed areas of the film.[56] However, organic based solvents were used as the developing medium for these resists, and since the molecular composition in the exposed and non-exposed film regions are nearly equivalent, these resists were characterized as exhibiting relatively low contrast and resolution. To overcome the organic solvent induced swelling and deformation of the negative tone pattern, researchers at IBM designed an epoxy-containing copolymer of hydroxystyrene and dicyclopentyloxy methacrylate, which, in the presence of a radiation induced acid would become crosslinked.[57] In this instance, the non-exposed film regions could be developed out in an aqueous base developer solution. Additional aqueous base developable epoxy-containing resist formulations were reported by Conely, who used certain classes of poly and monomeric epoxide compounds as crosslinkers for PHS based resins.[58]


Matrix resins
OH
x
Cresol novolac
x
OH
Poly(hydroxystyrene)
Photoacid generators
$R_nZ^+X^-$
Onium salts
CCl_3
Cl
Cl
DDT
Ar
N
CCl_3
N
N
CCl_3
s-triazine derivatives
Crosslinking agents
$ROCH_2$
N
CH_2OR
N
N
$ROCH_2$
N
N
N
CH_2OR
CH_2OR
CH_2OR
Melamine derivatives
(CH_2OR)
Benzyl alcohol derivatives

7.13 Chemical structures for the components in negative acting CA resists based on an acid catalyzed condensation polymerization radiation induced reaction pathway.

Condensation polymerization reaction pathway

Negative tone resist formation resulting from an acid catalyzed condensation polymerization route has proven to be a versatile approach for formulating these types of CA resists. Three-component CA resist materials have been investigated, and they consist of an acid generating species such as an onium salt or an s-triazine derivative, a novolac or PHS binder resin, and a multi-functional melamine or benzyl alcohol derivative crosslinking agent. Figure 7.13 provides the chemical structures for each of the components in these resists, which have found utility in photo, x-ray, and electron beam lithography-based processes.[59–62] Development occurs in an aqueous alkaline solution, and the materials also referred to as 'acid hardened resist' exhibit high dry etching resistance in oxygen and halogen based plasmas.

The postulated radiation induced reaction pathway consists of the exposure generating an acid, which, during the subsequent post-exposure bake step, protonates, for example, a melamine derivative (which contains multiple sites for crosslinking) and liberates an alcohol to produce a nitrogen stabilized carbonium ion. Alkylation of the binder resin occurs at either the phenolic oxygen (O-alkylation) or at a carbon on the aromatic ring (C-alkylation), which leads to the regeneration of the acid.[63]

Electrophilic aromatic substitution reaction pathway

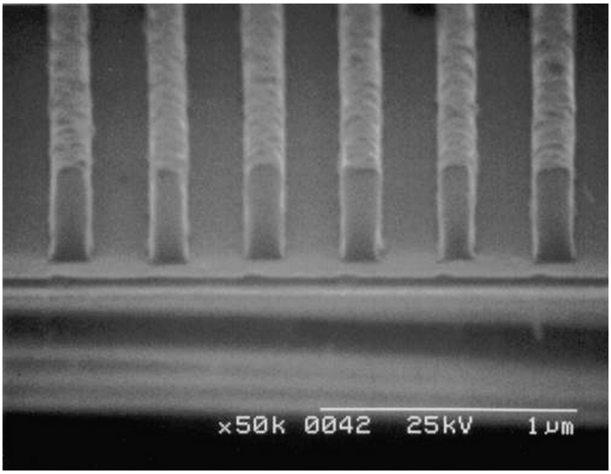
A third acid catalyzed reaction pathway used to generate negative tone patterns is based on the susceptibility of, for example, styrene based polymers to undergo an electrophilic aromatic substitution reaction.[64–66] The generalized radiation induced reaction pathway proceeds by the exposure generating the acid, which in turn during the subsequent post-exposure baking step reacts with a latent electrophile to form a carbocation intermediate. This intermediate then participates in an electrophilic reaction with the aromatic moieties of the matrix resin to form a crosslinked insoluble polymer network and regeneration of the acid. Latent electrophiles can exist in the resist in the form of an additive or as one of the monomers used to produce the resin material. Examples of latent electrophiles include dibenzyl acetate (additive) and copolymers of acetyloxymethylstyrene and 1,3-dioxane blocked benzaldehyde.[65,67]

Positive acting chemically amplified resists

The two most widely investigated acid catalyzed reaction pathways yielding positive tone pattern formation are based on deprotection and depolymerization chemistries.

Deprotection reaction pathway

As previously introduced in this section, single- and multi-component CA resist formulations have been reported to make use of the acid catalyzed deprotection reaction pathway to produce broad spectrum, high contrast, and resolution resist materials. The initial demonstration of this approach was reported by workers from IBM and is depicted in Fig. 7.9.[28] The two-component CA resist system set the foundation for the development of suitable 248 nm lithography positive acting resists. This predominant approach in the design of positive acting resist systems has been to take an inherently aqueous base soluble polymer and functionalize it with a 'protecting

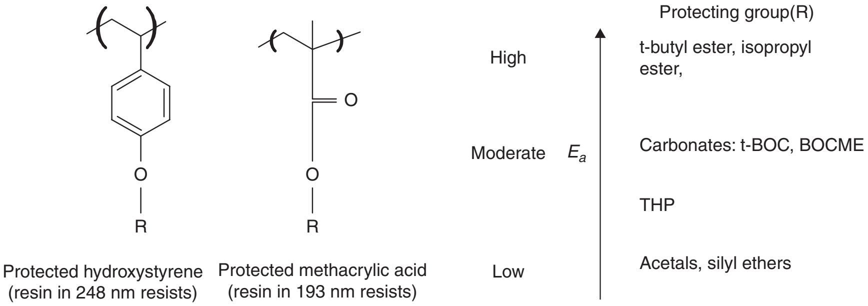


7.14 SEM micrograph of high fidelity coded 0.18 μm lines and spaces in a CA positive acting resist.

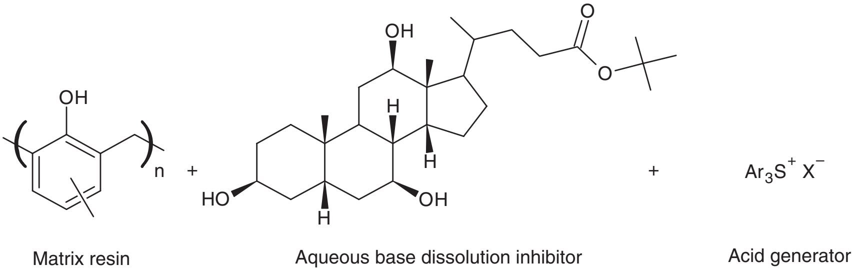
group' rendering it to be hydrophobic and only soluble in organic based solvents. An acid generating species is combined with the protected polymer, which upon exposure forms the acid, which either directly after the exposure or during the following post-baking step catalyzes the removal of the protecting group.[68] The removal of the polymer protecting groups renders the exposed regions of the film soluble in aqueous base developer and regenerates the acid. The exposure and bake-induced polarity change in the exposed vs non-exposed film regions resulted in these resists having characteristically high contrast (Gamma values as high as 10) and, as shown in Fig. 7.14, sub-micron patterning capability.[69,70]

Numerous families of functional groups were employed to protect the hydroxyl group of the polyhydroxystyrene resin, including t-butoxycarbonyloxy, tetrahydropyranyl, trimethylsilyl, and phenoxyethyl.[28,47,48,71,72] Likewise, alternative resins to the hydroxystyrene aqueous base soluble polymer, such as those based on poly(methacrylic acid) (PMAA), were also used. PMAA exhibited favorable optical and thermal stability properties[73] and found utility in mostly 193 nm lithography technologies. The generalized structure for both types of resins, and a selected list of protecting groups (R) classified according to the level of their activation energy (E_a) for the deprotection reaction, are provided in Fig. 7.15.

In addition, alternative approaches to the two component makeup of CA resist systems included multi-component resist formulations that removed the protecting functionality on the matrix resin and placed it either onto a molecular aqueous base dissolution inhibitor or the acid generating species.[31,74,75] An example of one such formulation is provided in Fig. 7.16. The


O
R
O
O
R
Protected hydroxystyrene
(resin in 248 nm resists)
Protected methacrylic acid
(resin in 193 nm resists)
High
Moderate
Low
E_a
Protecting group(R)
t-butyl ester, isopropyl ester,
Carbonates: t-BOC, BOCME
THP
Acetals, silyl ethers

7.15 Generalized chemical structures for protected polyhydroxystyrene and polymethacrylic acid used as the matrix resin for positive acting 248 and 193 nm CA resists and the corresponding types of protecting groups classified according to the level of E_a for the acid catalyzed deprotection reaction.

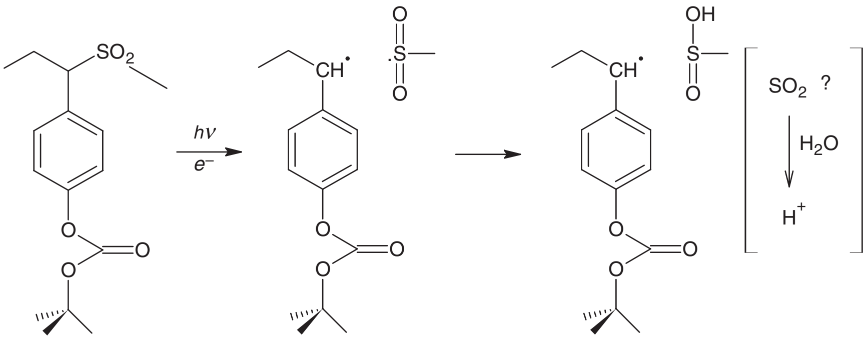

OH
n
+
OH
O
O
H
H
HO
H
OH
+
$Ar_3S^+ X^-$
Matrix resin
Aqueous base dissolution inhibitor
Acid generator

7.16 Chemical structure representation of a three-component tert-butylcholate based CA resist formulation.

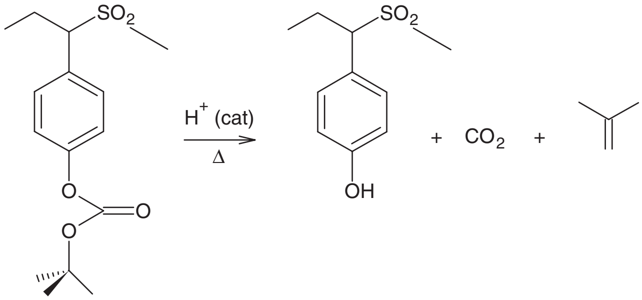
protected dissolution inhibitor (DI) shown in this example was chosen on the basis of its appreciable miscibility with hydroxyl based polymer matrix resins, high optical transparency at 248 nm, and having a large molecular volume, which provides for high solubility discrimination during development and image contrast.[75]

Conversely to this was the development of single component polymeric and molecular resists, where the acid generation and deprotection sites are all bound to the backbone of the polymer or on a single molecule.[76,6] For example, photo and electron beam sensitive, single component, aqueous base soluble positive acting CA resists consisting of copolymers of 4-tert-butoxycarbonyloxystyrene and sulfur dioxide (PTBSS) have been reported.[76] The exposure response and sensitivity of this resist was shown to be a function of copolymer composition, and independent of molecular weight. A schematic representation of the exposure induced reaction pathway in PTBSS is provided in the Fig. 7.17.

Radiation induced chemistry



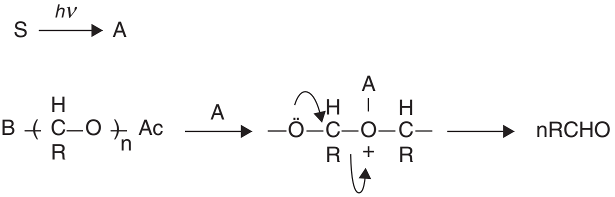
Acid catalyzed deprotection



7.17 Schematic representation for the radiation induced reaction mechanism occurring in PTBSS. In step one, polymer bound acidic moieties are formed *via* scission of the main chain of the copolymer. In step two, heat treatment of the resist film after exposure converts the copolymer to its aqueous base soluble poly (4-hydroxystyrene sulfone) form.

Depolymerization reaction pathway

A second radiation induced reaction pathway used to form positive tone patterns in CA resists is based on an acid catalyzed depolymerization reaction mechanism. This pathway was initially reported by Ito and Wilson at IBM, who made use of a two-component CA resist consisting of polyphthalaldehyde and an onium salt acid generating species.[77] Synthesis of aldehyde polymers proceeds at low temperature (<0°C) and the reactions are characterized as having a low ceiling temperature (T_c). Thermal stabilization (>100°C) of the polymer is accomplished by performing an alkylation or acylation end-capping reaction at the end of the low temperature polymerization reaction. Figure 7.18 provides the generalized acid catalyzed depolymerization reaction occurring in a polyaldehyde containing resist. The radiation generated



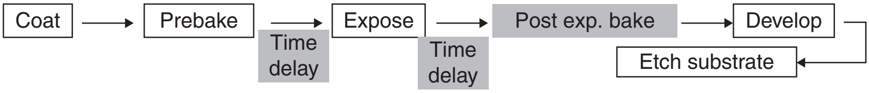
7.18 Two-step radiation induced chemical amplification reaction process yielding the dry development of a polyaldehyde resist sensitized with an acid generating species such as an onium salt.

acid is responsible for the catalytic cleavage of the acetal bond of the main polymer chain. This exposure process occurs above the T_c of the polymer, resulting in the spontaneous depolymerization of the polymer backbone and liberation of the starting monomer material. The radiation induced depolymerization process requires no subsequent thermal treatment or wet development step to form the positive tone pattern. Though the reduction of the necessary processing steps to produce the self-developing relief image is an attractive characteristic of these materials, liberation of the monomer during exposure can lead to appreciable contamination of the lithography tool. Subsequent investigation to eliminate the self-developing nature of poly(phthalaldehyde) was achieved through use of poly(4-chlorophthalaldehyde) and poly(4-trimethylsilylphthalaldehyde), which can be thermally developed during a subsequent post-exposure baking step.[78,79]

Follow-on use of acetals in three-component resist formulations was reported by researchers from Hoechst AG. They designed a resist in which the depolymerization pathway proceeded via acid catalyzed hydrolysis of O,O and N,O acetals.[80] Polymers having pendent silylether groups also represented another example of an acid catalyzed hydrolysis reaction leading to depolymerization of the starting material.[81] For this resist, radiation induced conversion to low molecular weight material occurs via cleavage of the Si–O bonds within the polymer structure. Silylether polymers in combination with a hydroxyl containing resin were added to the resist formulation to act as an aqueous base DI. Photo-exposure of this three-component resist generates an acid that catalyzes the hydrolysis and depolymerization of the silyl ether dissolution inhibitor, enabling the exposed film regions to develop using an aqueous alkaline solution.[82]

7.3 Chemically amplified resist process considerations

The introduction of CA resists into the DUV lithography pattern delineation process has been a key enabler for the continuation of the IC manufacturer



7.19 Patterning process flow diagram for CA resists requiring a PEB step. Highlighted boxes represent process conditions which are of critical importance to the performance of CA resists.

staying on the curve of Moore's law. CA resist performance requirements at the DUV lithography technology node such as high sensitivity, contrast, resolution, and dry etching resistance have been demonstrated. However, the introduction of use of this class of resist into the pattern formation process has led to the need to overcome potential limitations in achieving wide process latitude and the reproducibility required for establishing a robust resist patterning process.[83] The CA resist patterning process steps and associated variables that needed to be addressed and controlled are highlighted in Fig. 7.19.

In addition, early difficulties associated with introducing CA resists into a commercial IC production environment included:

- The observation for positive acting resists of a surface skin layer and non-vertical or T-topped shaped profiles resulting from incomplete development of the exposed film regions.[84]
- Understanding how the presence of the surface skin layer was dependent on the time from the post-coating bake to exposure, and from the exposure to the post-exposure bake step.[85]
- Identifying an acid precursor, the radiation generated acid, and the protecting group on the resin to be nonvolatile at the process set post-coating and post-exposure bake temperatures and during the exposure process.[86]
- The effect the radiation generated acid diffusing into the non-exposed regions of the film had on maintaining proper control of the resist CDs.[87]
- The effect the post-exposure baking conditions had on controlling the dimensions of the resist image and achieving the required ± CD control across the substrate.[88]
- For those resists that required a post-exposure baking step, hot plate units having ± 0.10°C temperature uniformity over its surface had to be manufactured.[89]
- The influence the composition of the surface of the substrate the resist was coated onto had on the resultant pattern profile.[83,84]

There continues to be a concerted effort by commercial IC manufacturers, academic and industrial R&D groups, and material and equipment suppliers to the IC industry to work collectively towards resolving the processing issues associated with CA resists. The following sections of the chapter will provide further insight into how these issues have been either brought under control or resolved so that these resists can be optimized to remain a key enabler to advancing 193 nm, EUV, charged particle, and the most forward looking lithographic technologies.

7.4 Chemically amplified resists for 193 nm lithography

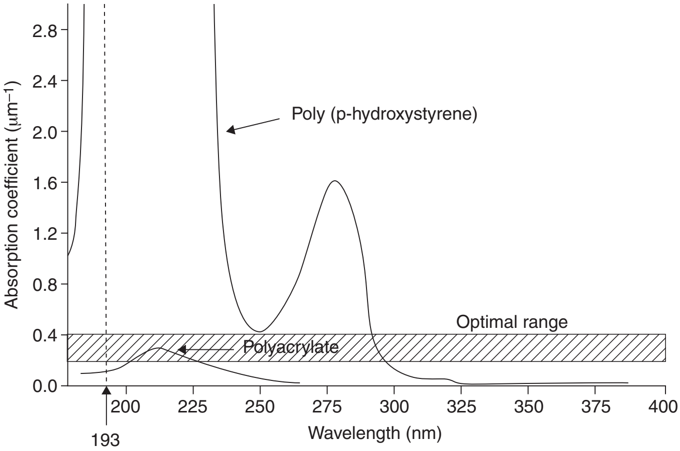
Given the extensive effort and development cost to commercialize resists based on acid catalyzed chemical amplification, the extension of their use beyond KrF lithography was paramount in the choice of the next generation or post 248 nm generation lithography technologies. Fabricating devices having minimum feature sizes below 0.18 μm led to the early development of small field 193 nm based lithography tools used for the development of the tool optics and resist materials.[90–92] In line with 248 nm lithography tools, the use of an Argon Fluoride (ArF) excimer laser centered at 193 nm[93] represented the continuation of the exposure source technology established at the longer wavelength technology node. The reader is also referred to Chapter 1 for a more comprehensive summary of this subject area.

The central search for an acceptable resist platform for the 193 nm lithography generation was to find a material having equivalent properties to the phenolic polymer and styrene–acrylate copolymer based systems used in the 248 nm generation resists.[94–96] Initial commercial 193 nm lithography systems, like their longer wavelength counterparts, performed the exposure in a gaseous environment above the wafer plane. The term 193 nm 'dry' lithography is used to describe this mode of exposure.[97,98] Investigation of the use of a liquid (water) coupling medium between the last lens optics element of the exposure tool and the surface of the resist having a refractive index > air, led to the development of a 193 nm 'immersion' (i) lithography system which greatly extended the life of 193 nm lithography generation.[99,100] More complex exposure strategies involving double- or multiple-pass exposures have extended the commercial use of this lithography technology even further, beyond all expectations.[101,102] Each of these exposure strategies imparted new and challenging issues to the development of a resist system compatible with these approaches. The following provides details of the history of the evolution of CA resist materials used throughout the extended life of 193 nm (ArF) lithography. As a starting point, the performance requirements that the resist must exhibit for each of the 193 nm exposure strategies are as follows:

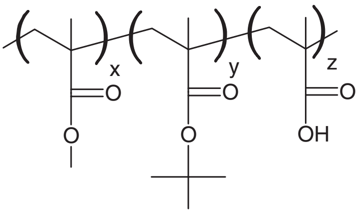
- Be a single layer process to minimize a new learning curve in the manufacturing environment and to keep material and fabrication costs as low as possible.

- Strong adhesion to the variety of substrate surfaces encountered in the fabrication process.
- Exhibit optical absorption properties affording high sensitivity and vertical pattern profiles. For a 0.40 μm thick resist film, an optical density of ~0.20–0.3 is required to enable a uniform exposure through the depth of the film.
- Thermal stability to enable the annealing of films above the T_g of the resist for ease of operation in the fabrication environment.
- <0.18 μm resolution for line and space pair pattern geometry.
- Process dose/photospeed equal to 10–50 mJ/cm^2.
- Pre- and post-exposure time delay latitude.
- ± 10% CD process latitude for <0.18 μm feature size.
- Minimal ΔCD vs post-exposure bake (PEB) conditions (<2 nm/°C).
- High dry etch resistance/low film loss in oxide (CF_4/O_2) and metal (Cl_2, BCl_3) plasma etch processes.
- Must use industry standard 0.26 N TMAH aqueous base developer.

Initial investigation of the optical properties (λ = 193 nm) of the phenolic polymer widely used in 248 nm resists revealed, as shown in Fig. 7.20, these materials to have unacceptably high absorption properties at this wavelength. Alternatively, as also shown in Fig. 7.20, polyacrylate based polymers do exhibit nearly transparent optical properties at 193 nm and represented viable candidates as resist materials for use in ArF laser lithography. IBM workers identified a 'version 1' acrylate-based, single layer, positive acting CA resist that could be patterned at both 248 and



7.20 UV absorption spectra of PHS and polyacrylate based polymers. Dotted line is used to compare the absorption values for the two materials at λ = 193 nm.



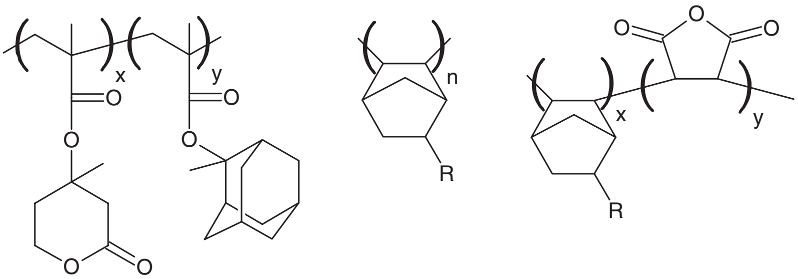
7.21 Chemical structure for the terpolymer present in the IBM 'version-1' 193 nm resist.

193 nm exposure wavelengths. This resist platform was highly useful in the evaluation of the earliest version 193 nm lithography tools.[103,104] The version 1 resist makes use of a terpolymer containing the monomers methyl methacrylate (MMA), t-butyl methacrylate (TBMA) and methacrylic acid (MAA).[105] The terpolymers are prepared using standard free radical initiated solution polymerization methods, and the reactivity ratios of the monomers affords considerable flexibility in preparing a broad range of terpolymer compositions. Figure 7.21 is a chemical structure representation of the MMA-TBMA-MAA terpolymer. The choice of each monomer, and their combinations in the terpolymer provided for optimizing the required functional properties of the resist. Good substrate adhesion, appreciable compatibility with the acid generating species in the film and a high T_g could be achieved with this resist.

Positive resist pattern formation occurs through an acid catalyzed deprotection reaction pathway involving the removal of the acid labile t-butyl group to form the resultant carboxylic acid.[106]

Two drawbacks to this formulation were: (i) the relatively high oxygen content in the resist resulted in low plasma dry etching resistance in a number of halogen dry etch recipes, and (ii) a very dilute aqueous alkaline solution (0.02 N) used in the development process precluded the possibility of using the industry standard 0.26N TMAH developer.

In the follow on generation of positive acting 193 nm CA resists, improvement in the plasma dry etch stability was addressed by incorporating alicylic group(s) into the polymer structure.[107,108] Early work by Onishi & Gokan[13,109] and then later on by Kunz[110] provided direction on the impact the molecular structure and composition of a material had on its dry etching resistance. The latter study concluded increasing the mass of the resist existing as carbon atoms present in a ring structure enhanced its durability in a dry etching environment. The alicyclic group represented an example of a carbon scaffold ring structure that could impart a substantial increase in the carbon content within the polymer. The lack of conjugation within the alicyclic (non-aromatic) groups yielded the additional attractive property of having high optical transparency at 193 nm. Alicyclic moieties have been introduced into the main

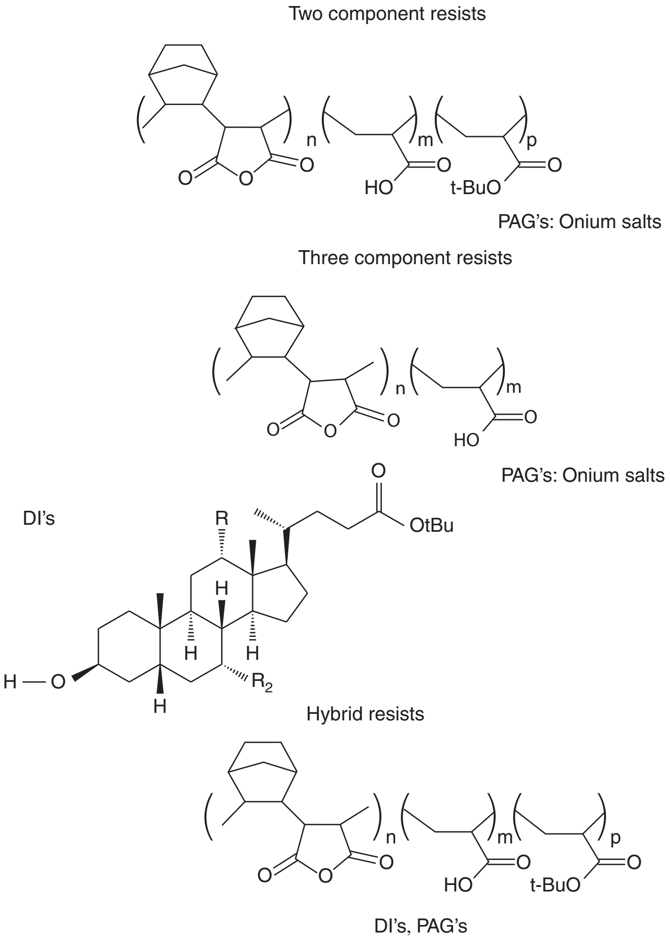


7.22 Example of chemical structures for alicyclic polymers incorporated into 193 nm resists.

chain of the polymer and as an inert side chain group where they may serve solely to improve stability in a plasma environment, or as an acid labile group in acrylic polymers.[111] Examples of alicyclic groups include those derived from adamantine, norbornene, and substituted norbornenes.[108,112] Structural representations of resist containing alicyclic groups are shown in Fig. 7.22.

Resists containing alicyclic groups were shown to improve stability in a plasma dry etch environment vs all acrylic based polymers.[107,113,114] High alicyclic content resists, however, are appreciably hydrophobic and difficulties arose in developing these materials in the standard aqueous alkaline developer. Co- and ter-polymerization with more polar groups, such as those containing a lactone functionality, improved the development characteristics of the resist. Addition of other methacrylate monomers afforded improved lithographic performance and compatibility with the 0.26 N TMAH developer solution.[115]

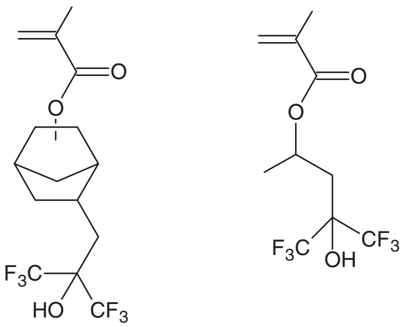
Another route to improving plasma dry etch stability of 193 nm resists was reported by researchers at Bell Laboratories,[116] who incorporated alicyclic moieties into the backbone of the polymer present in the resist formulation. Specifically, they made use of cyclic-olefin (norbornene) maleic anhydride alternating copolymers with inclusion of butyl acrylate and acrylic acid. Free radical solution polymerization of these alternating copolymers offered a controllable and scalable synthetic pathway. Use of acrylate vs methacrylate monomers in the feed mixture for the polymerization facilitated improving the yield of the reaction. These polymers exhibited high thermal stability ($T_g > 200°\mathrm{C}$), strong substrate adhesion, and solubility in an aqueous base solution.[117] The ability to use reduced amounts of carboxylic acid in the overall resist composition while maintaining aqueous base dissolution was a consequence of the hydrolysis products of the polymer anhydride group during the development process. This enhancement in development performance also minimized any swelling when using the industry standard TMAH aqueous base developer. Using this approach, a number of two or


Two component resists
n
m
p
O
O
O
HO
O
t-BuO
O
PAG's: Onium salts
Three component resists
n
m
O
O
O
HO
O
PAG's: Onium salts
DI's
O
R
OtBu
H
H
H
H—O
R2
H
Hybrid resists
n
m
p
O
O
O
HO
O
t-BuO
O
DI's, PAG's

7.23 Chemical structure representations for*: two-component resist* formulations based on a terpolymer of norbornene-maleic-anhydride-acrylic acid and-t-butyl acrylate with an acid generating species; *three-component resist* formulations which include a cholate ester DIs, acid generating species and acidic base resin and *hybrid resist* formulations based on these two or more component resist formulations.

more component resists based on the cyclic-olefin backbone polymer resist formulations were devised and, as provided in Fig. 7.23, optimized for use in 193 nm lithography.[118]

Concurrent with the development of 193 nm resists was the investigation of a lithography tool generation using a fluorine (F_2) excimer laser



7.24 Chemical structure of acrylate-based monomers containing the hexafluoroisopropanol group which have been incorporated into 193 nm resist formulations.

having an emission at 157 nm.[119] In the case of 157 nm lithography, incorporation of functional groups containing fluorine such as the fluoroalcohol moiety provided a pathway to identifying optically transparent materials at this wavelength.[120] The fluoroalcohol group is also transparent at 193 nm, and this property, in combination with the need to control the size of the CD during the PEB step, led to these materials also acting as high performance 193 nm resists.[121] One approach to including fluorine as part of the resist consisted of incorporating the hexafluoroisopropylalcohol (HFIPA) group into the side chain of methacrylate-based resists.[122,123] Figure 7.24 provides examples of methacrylate monomers containing the pendent HFIPA group.

Incorporation of the fluoroalcohol group into the side chain of methacrylate resists not only improved the optical transparency at $\lambda = 193$ nm, but also has been used to control the resist's aqueous base dissolution properties. The group's acidity is similar to that of a phenol[124] and, through a balance with other components in the resist, can be used to regulate the influence of the presence of the carboxylic acid groups on the dissolution behavior of the material.

In addition, this group, when present in the resist, acts to reduce the sensitivity to the conditions used in the PEB step (ΔCD/°C).

7.4.1 Chemically amplified resists for 193 nm immersion (i) lithography

The extension of using 193 nm light as the exposure source to fabricate device size features well below 100 nm was realized through the introduction of commercially available 193 nm water immersion lithography tools.[125] Immersion lithography provides an imaging system with an effective numerical aperture >1 and an increase in depth of focus (DOF).[126] The key to the success of this lithography system was to make use of the commercially available 193 nm CA resists with the immersion exposure process. This would provide for the rapid continuation of 193 nm lithography and accelerate closure on the need for continued development of the exposure tool and resists for 157 nm lithography.

The choice of water as the immersion liquid was driven by it having a refractive index at 193 nm equal to 1.44 and an extremely low absorption coefficient on the order of 0.04 cm^{-1}.[127] Ultra-high purity water is required, and the critical challenges associated with having the surface of the resist film coming into contact with the water immersion liquid include:

- Extraction (dissolution) of any of the components of the resist film into the water medium.
- The degree to which water enters (permeates) into the resist film.
- The observation of any outgassing of the resist film during exposure resulting in the presence of bubbles in the liquid medium.
- The presence of any particulates formed during the resist–water interaction.
- The durability of the resist film to remain intact/not degrade during the high speed exposure scan rate process of the lithography tool.

Collectively, the observation of any of these processes will adversely affect the overall lithographic performance of the 193 nm CA resist. Strategies to mitigate the influence of the water immersion exposure process on resist performance include:

- Use of a top coat protective film coated onto the surface of the resist film.
- Introduction of additives into the resist, which enables the resist to perform to specification without the need of a top coat protective film.

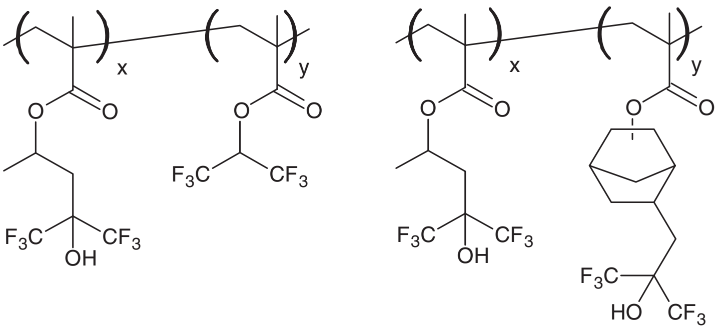
Top coat protective films

The function of the top coat film is: (i) to prevent the leaching or entering of components of the resist film into the water immersion medium, and (ii) to mitigate any added defects throughout the lithographic patterning process. The top coat, if functioning properly, can therefore eliminate any variability in the composition of the resist film during exposure and protect the illumination lens system of the lithography tool and the underlying layers of the resist-coated substrate from deposition of any exposure generated contaminants. Early version top coat films were only soluble in organic solvents.[128] This necessitated the introduction of an additional processing step to remove the top coat film prior to development of the exposed resist film in the standard aqueous alkaline developer. Later generation top coat films consisted of materials that were designed to be soluble in the standard aqueous alkaline resist developer and eliminated the need for introducing additional post-exposure processing steps.[129] However, for either type of top coat material, the presence of it on the resist film introduced new variables that could affect the reproducibility of the exposure process and performance of the resist and these include:

- The degree of penetration of water into the top coat layer and into the resist film below it.
- The wettability by the water onto the surface of top coat film during the scanning exposure tool process.
- The presence of an interfacial layer resulting from the mixing of the two films during the coating of the protective film onto the resist surface.
- The promotion of any contaminants and the control of the number of defects resulting from the intimate contact the water layer has during the wafer exposure process and during the removal in the resist development step.

The cost of the top coat material and the need to not add a processing step to remove the film greatly influenced its choice. Elimination of the added processing step could be achieved using a top coat material that was alkali soluble and would readily wet the surface of the resist film which it is spin coated onto.[130] A number of top coat films containing materials having fluorocarbon, hydrocarbon and organosilicon functional groups were shown to increase the water contact angle values.[131] The resultant reduction in the surface energy at the water-top coat film interface also has the effect of eliminating the mechanical failure or removal of either film layer from the Si substrate and the likelihood of introducing defects into the device being fabricated.[132] Acrylic based polymer materials containing fluoroalcohol functional groups, which were initially investigated for use as positive acting CA resists for 157 nm lithography and early version 193 nm resists, found utility as top coat materials.[133] These films are highly transmissive at 193 nm, appreciably wet many commercial ArF resist films, and are prepared using a low cost manufacturing process.[134] Figure 7.25 provides examples of top coat materials containing the fluoroalcohol groups.

Fine control in meeting the desired water contact angle and solubility in aqueous alkaline solution was achieved in these materials by optimizing the structure of the fluoroaliphatic side chain group.[135] Use of hydrocarbon or ether based spin casting solvents for the top coat material aided in minimizing the presence of an interfacial layer between the resist and top coat films, thereby enhancing their use as a top coat film.



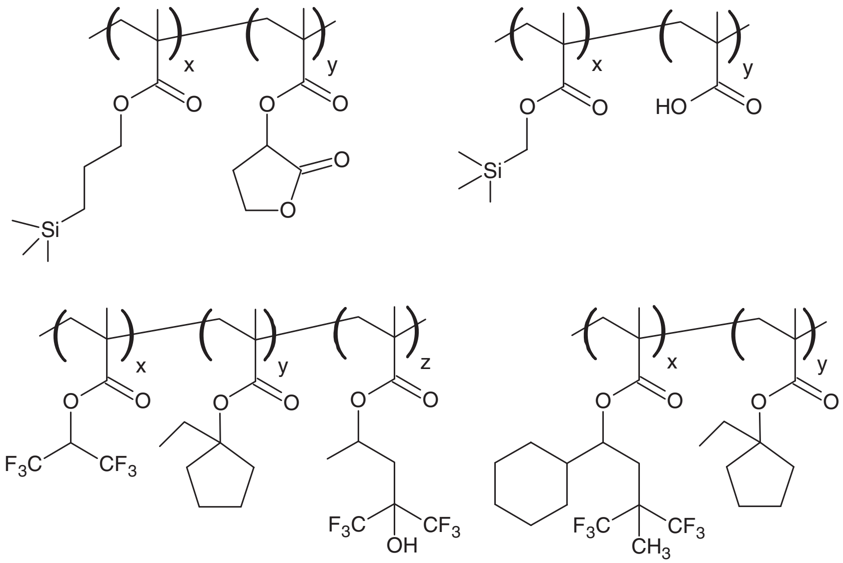
7.25 Chemical structure representation of fluorine-containing copolymer materials functioning as top coats for use in 193 nm immersion lithography.

Top coat films can also be designed to be an effective antireflective coating (ARC) layer.[136] The appropriate thickness and optical properties required to act as an ARC layer must be balanced by maintaining the key film requirement of protecting the resist film from deleterious interactions with the water immersion layer. Top coat materials have been used to improve the observed quality of the resist image after development. This is made possible by incorporating an acidic functionality into the top coat material. This has the effect of reducing the presence of a post-development surface skin layer of the 193 nm resist film by compensating for any unintended reduction in acid concentration at the resist surface after it has been exposed.

Minimizing defect density resulting from the use of a top coat film represents one of the most challenging requirements for use in 193i lithography. The pertinent processes to reduce the observed defects are to eliminate the pulling of the resist film during the scanned exposure step and the presence of any remaining water droplets after the resist film has been exposed and continued onto the post-exposure processing steps.[137]

Immersion compatible/top coat free resists

CA resists that do not require a top protective coating to be exposed in a 193 nm immersion lithography tool represent a more attractive approach to reduce process complexity, defectivity, and cost. One path to achieving this is to reduce the hydrophilicity of the resist by incorporating, for example, a hydrophobic fluorinated alicyclic-based monomer into the resist.[138] In addition, there is the need to design the acid precursor and radiation generated acid product species to not dissolve or leach into the immersion liquid. This requirement represented a formable challenge in identifying a top coat free resist material.[139]



7.26 Chemical structure of silicon and fluorine containing methacrylate co- and ter- polymers used as additives in a top coat free 193 nm immersion lithography resist process.

A more prevalent path has been to incorporate an additive into the resist that functions to produce an *in situ* topcoat.[135] The additive operates by having a lower surface energy than the other components in the resist formulation, which enables the additive to rise and become the predominant species at the surface of the resist film during the spin coating process. The high concentration of the additive at the surface effectively provides the equivalent functions of a top coat film previously described and to critically provide the proper contact angle for minimizing the generation of defects during the scanned immersion exposure tool process. Additives that have been investigated include organosilicon- and fluorine-containing polymers. The low surface energy groups are incorporated into the additive as either part of the backbone structure or as part of the side chain of the base polymer making up the material. A polymethacrylate backbone has been widely used, and examples of silicon- and fluorine-containing additives are depicted in Fig. 7.26.[140,141]

A second consideration in the identification of the additive is, it does not inhibit the exposed areas of the resist film from dissolving in the aqueous base developer solution. Insolubility in the developer solution can lead to an increase in the roughness of the patterned resist image and the number of defects in the final device structure. Solubility of the additive in the aqueous base developer solution is achieved by incorporating non-polar acid labile groups into its polymeric structure, such as that also shown in Fig. 7.26. The exposure and PEB steps simultaneously drive the acid catalyzed deprotection reaction of the additive and resist resin components to ensure complete dissolution in the developing solution. The use of a top coat free CA resist has cleared the way for returning to employing a cost effective, low defect level, single layer resist in the 193 nm immersion lithography process.

7.4.2 Resists for double patterning, double exposure 193 nm lithography strategies

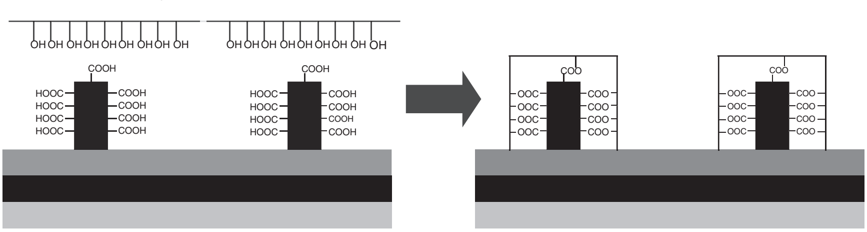
The thrust of the IC manufacturing community to further push the resolution envelope of 193 nm immersion lithography for use in the 22 nm technology node generation has led to the need to develop more complex and challenging exposure and patterning processes.[142] In order to press on to finding a solution for the 22 nm technology node, a number of resolution enhancement techniques have been proposed that do not radically alter the lithography process used in today's commercial IC manufacturing world.[102] With the exception of the '*Self-Aligned Spacer Double Patterning Process*', which is discussed in a latter section, the heart of these resolution enhancement techniques is the requirement to perform a minimum of two exposure steps per wafer level. The double exposure process enables the pattern of a complex photomask to be broken up into two individual exposures, thereby relaxing the difficulties associated with the single-exposure printing of the entire layer. The two exposure processes are in effect combined to print the pattern of the original mask. This approach relaxes operating the fabrication process at the lower limit of k_1, and enables realistic implementation into a manufacturing environment. Distinctions or variations in the two exposure step process have opened up a frontier for the resist chemist to develop new materials and processes to meet the 22 nm device design rule requirement.[143,144] The following sections highlight the various two exposure step processes being pursued with 193i lithography and are also further discussed in the section in Chapter 1 entitled 'Multiple Patterning Lithography.'

Double patterning lithography (*DPL*)

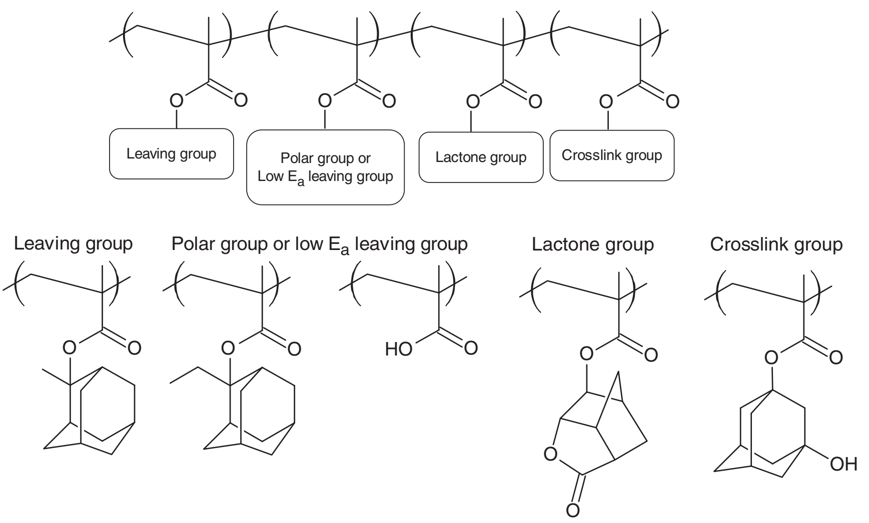
Of the most recent resolution enhancement techniques, the double patterning, single level process represents the quickest approach to turn this on in production. Many variants of this technique have been proposed and practiced. These include, but are not limited to:

- *Litho-etch, litho-etch* (*LELE*) *process*. Two complete lithography and subsequent etch processes are performed to fully pattern a single device layer.[145] In this process, two resists and two sequential processes consisting of resist coat–bake–expose–bake–develop–etch are performed. In this instance, the wafer must be removed from the lithography system and brought back to be realigned, and then exposed a second time. The two complete lithography steps involving the resist can make use of the same or a different material, and can use either a negative or positive acting resist. The straightforwardness of this approach, and the lack of the situation where the first resist layer is coated onto an existing patterned resist layer, provided for rapid acceptance of this resolution enhancement technique.

- *Litho-litho-etch* (*LLE*) *process.* With this, the requirement of performing a second etch step is removed, thereby simplifying the overall complexity of fabricating the final device pattern.[146] However, this process requires the second resist coating to be performed directly onto the initial patterned resist layer. There is the concern related to the initial resist patterned layer intermixing during the second resist coating process. This has the potential to distort the images formed from the coating of each resist layer during subsequent baking and development steps. The following processes have been used to overcome the potential for intermixing of the two resist layers:
- *Litho-freeze-litho-etch* (*LFLE*) *process*. To address the resist intermixing issue, two processes to '*freeze*' the first formed resist pattern have been investigated.[147] In the '*chemical freeze*' process, a distinct material is introduced into the process prior to the coating of the second resist coating step. The chemical agent is dispensed onto the existing resist pattern and induces the patterned resist material to become highly crosslinked. The crosslinked resist is now rendered insoluble in the spin casting solvent used in the second resist coat, thereby greatly reducing or eliminating any intermixing of the resist layers. Figure 7.27 provides a schematic representation as to how crosslinking occurs in the '*chemical freeze*' process.



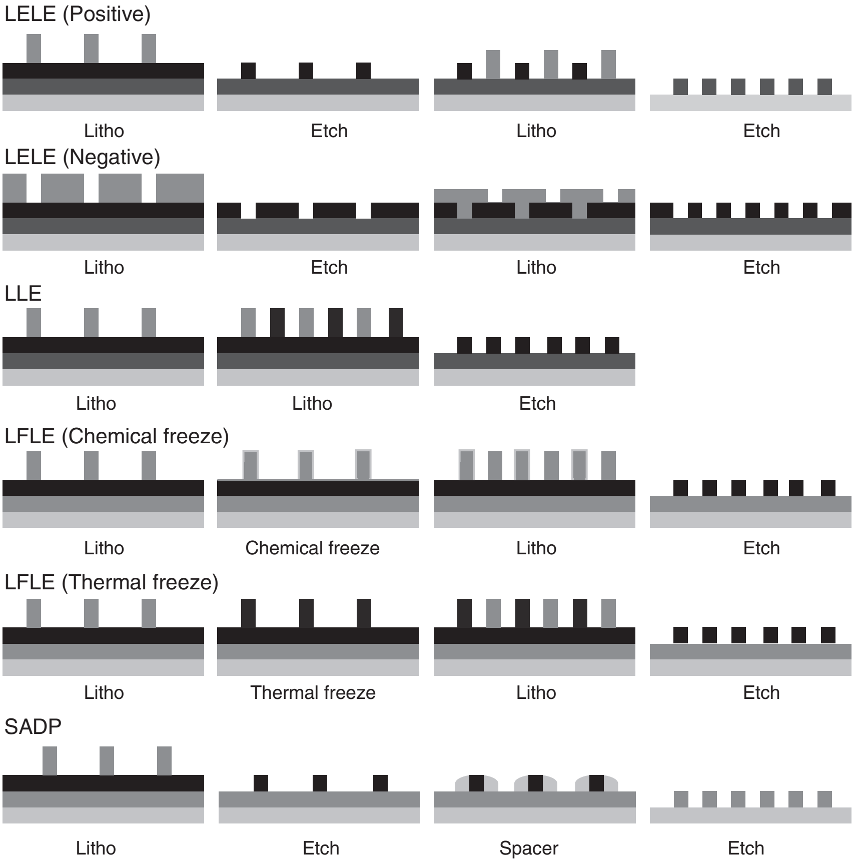
7.27 Diagrammatic representation of the '*Litho-chemical freeze-litho-etch*' process in which crosslinking of the initially patterned resist image occurs prior to second resist coating step.



7.28 Chemical structure and functional requirements of the monomers comprising the resist involved in a 'litho-thermal freeze-litho-etch process'.

Alternatively, instead of using a chemical crosslinking agent, the first patterned resist layer is thermally treated (heated) to induce crosslinking within the material. The '*thermal freeze*' process has the advantage of not introducing any additional chemicals into the overall patterning process and conforms to using the widely accepted lithography process baking step. Figure 7.28 provides the framework chemical structure and functional requirements for the polymers comprising a resist that is involved in a '*thermal freeze*' process.

- *Self-aligned spacer double patterning process* (*SADP*). This process has been the most heavily exercised in the 193 nm immersion lithography printing of a semi-dense pitch core grating pattern, which typically uses a 1:3 line-to-space pattern arrangement.[148] As shown in Fig. 7.29, after the resist is patterned, the remaining structures are coated with an inorganic SiO_x or Si_xN_y layer *via* a chemical vapor deposition or spin coating process to form spacers on the resist side wall. The spacers function to split the pitch of the existing pattern by way of the top coated part of the resist pattern being subjected to a polishing process. This results in selectively removing the inorganic layer from the resist surface, but does not affect the material that has been deposited onto the side wall of the resist images. The remaining uncoated part of the resist pattern is then removed during the subsequent dry/plasma etching step, leaving only the standing inorganic material side wall structure to act as the etch


LELE (Positive)
Litho
Etch
Litho
Etch
LELE (Negative)
Litho
Etch
Litho
Etch
LLE
Litho
Litho
Etch
LFLE (Chemical freeze)
Litho
Chemical freeze
Litho
Etch
LFLE (Thermal freeze)
Litho
Thermal freeze
Litho
Etch
SADP
Litho
Etch
Spacer
Etch

7.29 LELE, LLE, LFLE and SADP process flow diagrams.

mask during this pattern transfer step. This process produces in the end a higher density, smaller 1:1 pitch grating structure. In certain instances, additional lithography processing is required whereby a separate photomask or 'cut mask' is used to trim the ends of the lines that were deposited with the inorganic material. This approach provides for superior overlay performance over other DEP processes, and is most suitable for forming regularly arranged dense patterns.

Double exposure lithography (DEL)

DEL, unlike the previously described resolution enhancement techniques, does not require the coating of two resist layers, deposition of an additional material, or the removal of the substrate from the lithography tool during any part of an individual device level patterning process.[149] In DEL, a single layer of resist is coated onto the substrate and is subjected to two sequential exposures followed by one develop and one etch step. The exposures are positioned to be staggered with respect to each other, and are required to have twice the desired pitch in order to produce the final pattern. This approach does require increasing the number of masks to two for each device level pattern, but substantially reduces the overall number of processing steps relative to the other resolution enhancement techniques. DEL, however, requires the resist to have a nonlinear exposure response and this precludes the use of the existing platform of resists used in 193i lithography. This is due to what is referred as the 'memory effect' occurring between the two exposures performed in the single layer of resist.[150] Various approaches have been described by C.G. Willson *et al.*[149] to identify a material with a nonlinear exposure response. These include contrast enhancement layers (CEL), two-photon materials, intermediate state two-photon (ISTP) materials, and optical threshold layers (OTL).

7.5 Resists for extreme ultraviolet lithography (EUVL)

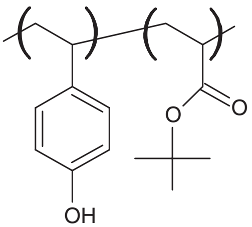
The following section provides a brief summary of EUVL resist materials based on their polymer platforms.

7.5.1 Methacrylate-based polymer platform

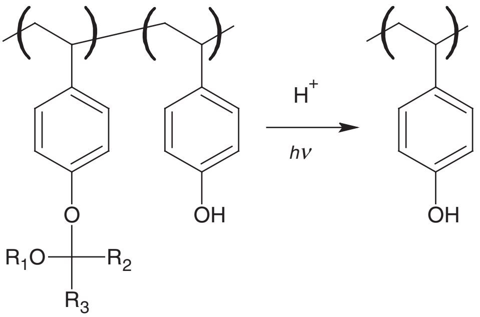
Polymethylmethacrylate (PMMA), a well-known e-beam resist[151], has also been investigated for use in EUVL.[152,153] PMMA is a non-CA resist and, equivalent to what occurs when exposed to charge particle radiation, undergoes scission of the main chain of polymer when subjected to extreme ultraviolet (EUV) radiation. The scission events lead to a reduction in the molecular weight of the PMMA present in the exposed film regions and enables selective development of those film areas in an organic based developer.[154] PMMA has demonstrated low line edge roughness (LER) and high resolution performance when exposed by an EUV interference lithography system. Half-pitch grating patterns having 20–50 nm dimensions were delineated using this approach.[155]

7.5.2 Hydroxystyrene-based polymer platform

Environmentally stable chemically amplified photoresist (ESCAP) was invented at IBM as a DUV CA resist and was first explored for application in EUVL in 1999.[156] The polymer component in ESCAP is poly (t-butyl acrylate-co-p-hydroxystyrene) and the structure is shown in Fig. 7.30.



7.30 Chemical structure of the polymer/resin component in ESCAP resist.

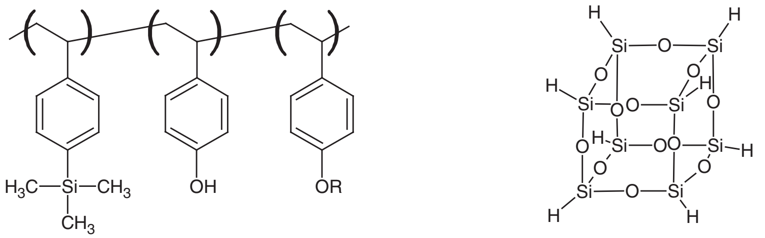


7.31 Chemical structure and chemical amplification reaction process in KRS resist.

Chemical amplification in ESCAP is based on the deprotection reaction mechanism pathway, and showed high photosensitivity, good stability in a dry etching environment, and compatibility with a PEB aqueous base development process. 50 nm dense and 30 nm semi-dense lines in 80 nm thick resist could be achieved using the ESCAP version XP9947W resist when exposed in a EUV interference lithography tool.[157]

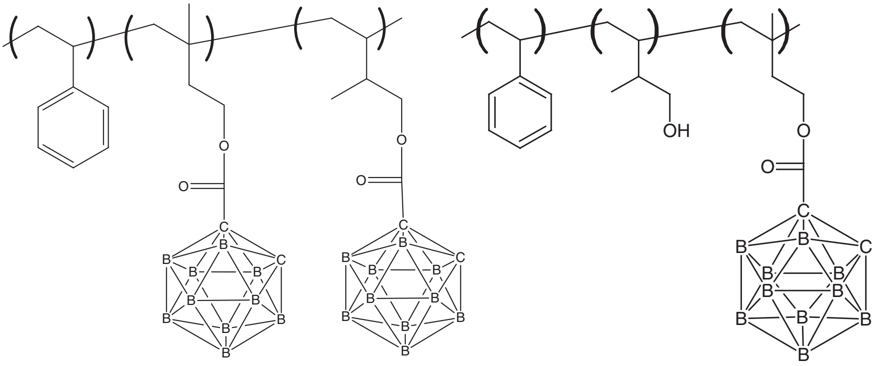
7.5.3 Ketal protected hydroxystyrene-based polymer platform

KRS resist is a CA positive acting resist based on a ketal protected poly(p-hydroxystyrene) platform.[158] The low activation energy for deprotection of the ketal protecting group enables this process to occur at room temperature, and is shown in Fig. 7.31. This room-temperature reaction has the advantage of reducing resist image blur when exposed to EUV and other forms of radiation. KRS resist has demonstrated the ability to resolve 25–45 nm dense line patterns when exposed using LBNL's 0.3 NA EUV micro-field tool.[159]



Poly(trimethylsilylstyrene-co-p-hydroxystyrene)

Hydrogen silsesquioxane (HSQ)



Boron containing resist polymers

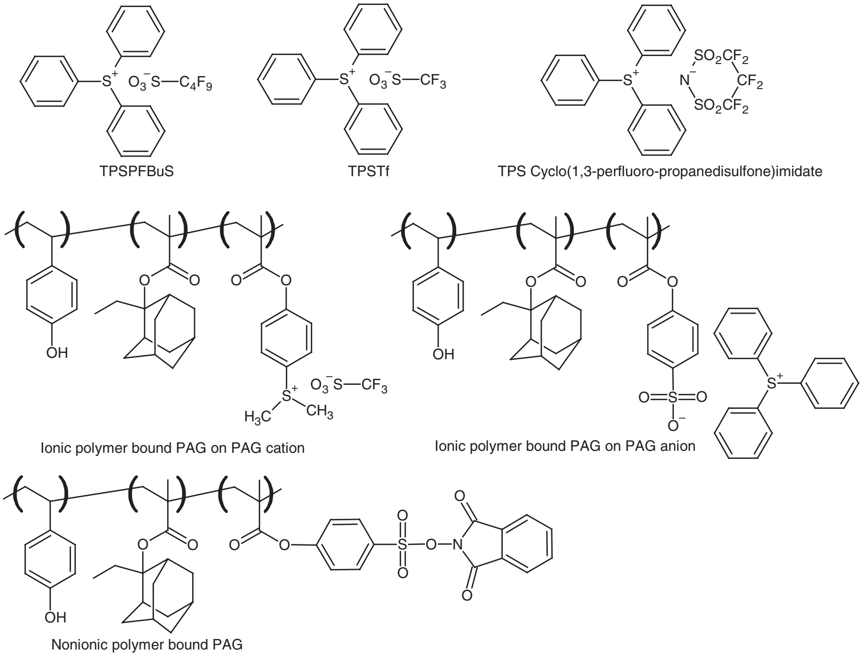
7.32 Chemical structure for silicon containing resists [poly (trimethylsilylstyrene-co-p-hydroxystyrene), HSQ and boron containing EUVL resist polymers.

7.5.4 Low EUV atomic absorption cross-section resists

Transparency to EUV radiation is a critical requirement in designing a resist for this lithography technology. Incorporation of elements such as H, C, B, Be, and Si into the resist, which have low atomic absorption cross-section coefficients at $\lambda = 13.5$ nm is helpful in meeting the transparency requirement. Figure 7.32 provides the chemical structure for exemplary silicon and boron containing EUVL resists.[160,161]

7.5.5 EUVL acid generating species

Acid generating species used in earlier generation 248 and 193 nm CA resists are also commonly used in EUVL resist formulations. Figure 7.33 illustrates the structure for a few exemplary acid generating species used in EUVL resist formulations.[162–164]


TPSPFBuS
TPSTf
TPS Cyclo(1,3-perfluoro-propanedisulfone)imidate
Ionic polymer bound PAG on PAG cation
Ionic polymer bound PAG on PAG anion
Nonionic polymer bound PAG

7.33 Chemical structure descriptions of acid generating species incorporated into CA EUVL resists.

One innovative structural design[164] used in a select number of these acid generating species is the moiety responsible for the exposure induced acid generation to be chemically attached onto the side chain of the polymer present in the resist. This has the desirable effect of minimizing resist image blur, and reducing the diffusion and level of outgassing from the resist film. These attributes collectively enhance resist resolution, and reduce LER without sacrificing sensitivity to EUV radiation.

7.6 Resists for electron beam lithography

Concurrent with the development and advancement of optical lithography has been the need to establish a commercialization path for the manufacture of the photomask or reticle with the highest level of perfection. The requirement to produce a ‘perfect’ mask (which is the focus of Chapter 5), whereby the computer aided design (CAD) pattern of a device layer is exactly replicated onto the mask, has also led to the development, evolution, and commercialization of direct write electron beam lithography systems (DWEBLS). Chapter 3 provides a comprehensive description for these exposure systems. In addition to the use of these exposure tools to fabricate photomasks, DWEBLS has also found utility as a method to directly form the pattern in a resist layer coated onto a semiconductor substrate, thereby bypassing the need to first fabricate a photomask. Direct write electron beam is the lithography of choice to carry out the fundamental research required to understand how new materials and process approaches can be used to fabricate advanced devices. However, use of DWEBLS for high volume manufacturing of ICs is challenged by the serial writing nature of the exposure tool. High throughput approaches that have been pursed include the development of multi-electron beam exposure systems and projection-electron lithography (PEL) tools.[165]

In addition to making use of electrons as the exposure source, a beam of ions such as Gallium, and Helium have been used in DW exposure systems. A more detailed description of focused ion beam lithography systems can be found in Chapter 4.

The development of suitable charge particle resist materials in particular for the photomask industry required the need to identify both positive and negative acting resists. The typical Chromium-on-glass mask fabrication process is analogous to what is displayed in Fig. 7.2. Exceptions are that the exposure source is electron based, which as stated is a direct write process requiring no mask fixture above the plane of the resist layer during exposure. The choice of the resist tone in the mask fabrication process is driven by minimizing the required exposure area and corresponding time for a given pattern layout.[166] The selection process is summarized as follows:

> **Selection of the tone of resist depends on the ratio of the clear/opaque area of the pattern.**
> **Clear area < Opaque area – use positive resist**
> **Clear area > Opaque area – use negative resist**

Reproducibility and relative simplicity in the photomask fabrication process stipulated that the photomask blank supplier provide the substrate with a stable/long shelf life resist film coated onto it and ready for exposure. The need to transfer the resist pattern into the Chromium photomask layer required the resist to withstand wet and dry etching environments unique to that layer.[167] Table 7.3 summarizes the resist material properties requirements when fabricating a standard binary Cr-on-quartz photo mask. Based on these required properties, the early development of resist materials for use in the mask fabrication took a different path from what is used in optical lithography.

Electron beam resist sensitivity is of critical importance in the manufacturing process from the standpoint of minimizing the exposure time of the mask and any heating of the quartz substrate during exposure. Exposure induced substrate heating and subsequent expansion can lead to linewidth variation and pattern placement errors on the fully fabricated mask.[168,169] In contrast,

Table 7.3 Cr-on-glass photomask electron beam resist property vs performance requirements

Property	Performance requirement
Solution shelf life	>6 months
Batch to bath reproducibility	<5% variation in composition and molecular weight
Film coating on photomask blank shelf life	>3 months
Thermal properties	Glass transition temperature $T_g > 80°C$
Wet etch Cr chemistry	No degradation or adhesion failure
Dry etch Cr chemistry	Min 1:1 selectivity in Cl_2-O_2 based plasma
Solubility	In environmentally safe spin coating and developing solvents
Stripability	Removal in commercial based stripping solutions or O_2, halogen-based plasmas

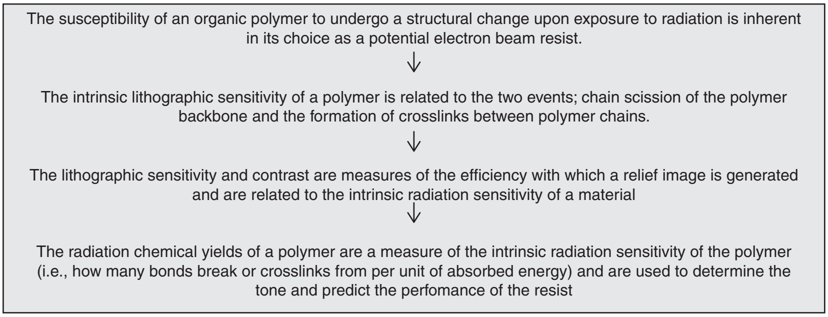
when a DWEBLS is principally used for nanofabrication purposes, resolution capability is the ultimate driver in the choice of the resist.[9] Finally, when multi or projection electron beam based systems are considered for potential use in high volume IC fabrication; the concurrent requirements of high sensitivity and resolution must both be considered in choosing a suitable resist.[170–172] Note that the reported resist sensitivities used throughout this section of the chapter are functions of the electron beam exposure tool accelerating potential setting, and are approximately inversely proportional to the beam energy.

7.6.1 Intrinsic property requirements of polymers as electron beam sensitive resists

Earliest version electron beam resists were single component polymers, which, upon exposure, induced fracturing of the main chain of the polymer or led to the formation of inter-chain crosslinks. Figure 7.34 provides the guidance and fundamental analysis approach used to determine whether the polymer will predominantly undergo main chain scission or inter-chain crosslinking.

G-values and their influence on tone/choice of resist

For resist applications, one can make use of the *G*-values to determine which radiation induced polymer degradation path predominates.[173] The radiation chemical yields are expressed in terms of *G*-values. *G* (scission),



The susceptibility of an organic polymer to undergo a structural change upon exposure to radiation is inherent in its choice as a potential electron beam resist.

↓

The intrinsic lithographic sensitivity of a polymer is related to the two events; chain scission of the polymer backbone and the formation of crosslinks between polymer chains.

↓

The lithographic sensitivity and contrast are measures of the efficiency with which a relief image is generated and are related to the intrinsic radiation sensitivity of a material

↓

The radiation chemical yields of a polymer are a measure of the intrinsic radiation sensitivity of the polymer (i.e., how many bonds break or crosslinks from per unit of absorbed energy) and are used to determine the tone and predict the perfomance of the resist

7.34 Step wise fundamental analysis approach to determine the predominant e-beam induced reaction process in a candidate resist material.

$G(s)$, equals the number of main chain scissions produced per 100 eV of energy absorbed, and G (crosslinking), $G(x)$, equals the number of crosslinks formed per 100 eV absorbed. The G-value is a structure-dependent constant similar to quantum efficiency in photochemistry.

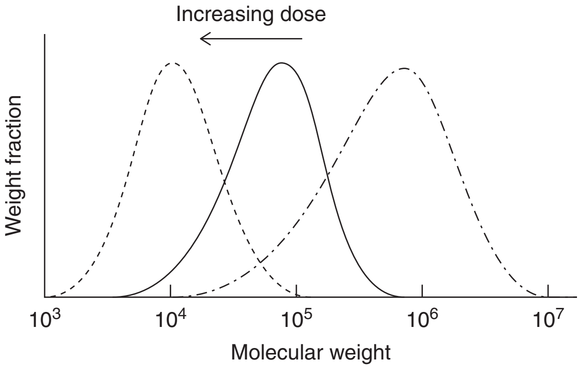
Specifically one can use the ratio of $G(s)/G(x)$ to determine the tone of the resist whereby:

when the ratio of $G(s)/G(x) > 4$, main chain scission reactions predominate, resulting in a decrease in the polymer molecular weight of the exposed vs the unexposed regions of the film. The difference in the wet development rate between the exposed vs non-exposed regions is then used to form a positive tone image.

when the ratio of $G(s)/G(x) < 4$, inter-chain crosslinking reactions predominate, resulting in an increase in molecular weight leading to an insoluble gel and negative tone formation.

G-values and the ratio of $G(s)/G(x)$ can be correlated to the contrast of the resist. In general, for positive resists a higher value of $G(s)$ and a lower value of $G(x)$ yield a higher ratio value, resulting in higher resist, contrast, and resolution capability. For negative resists the converse is true, whereby a higher value $G(x)$ and a lower value of $G(s)$ yield a lower ratio value, resulting in higher resist, contrast, and resolution capability. Best case scenarios are: for + tone resists $G(x)$ is zero; and for – tone resists $G(s)$ is zero. The worst case scenario is when $G(s)/G(x)$ approaches four. Competing scission and crosslinking reactions lead to a very low contrast resist. In polymer systems it is common to have both scission and crosslinking reactions occurring, and it is the threshold absorbed dose for one vs the other reaction pathway that defines the tone of the resist.

In the absence of a defined lithographic process for a polymer being investigated as a suitable resist, gamma (γ) radiolysis using a Cobalt-60 source of solid samples can be used to determine the G-values for that material.[174]



7.35 Change in molecular weight as measured by gel permeation chromatography vs e-beam exposure dose for polymethylmethacrylate (PMMA) positive acting resist.

The determination of the *G*-values requires measuring the change in the molecular weight of polymer as a function of the absorbed dose. Molecular weight changes, as a function of dose, can be measured using size exclusion chromatography. The incident dose of an electron beam lithography tool required to impart the absorbed dose into the polymer can be calculated knowing the composition of the polymer and exposure conditions of the electron beam system.[175]

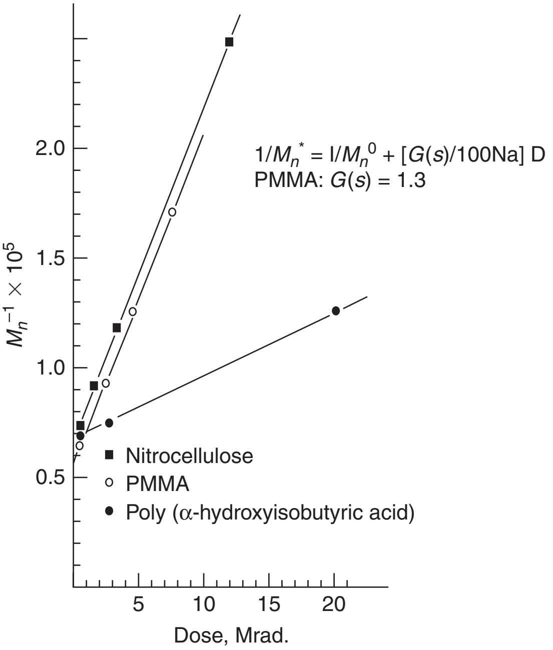
7.6.2 Positive acting electron beam resists

Figure 7.35 provides for the change in molecular weight vs electron beam exposure dose for polymethylmethacrylate (PMMA). Evidence of a decrease in molecular weight as a function of increasing electron beam exposure dose indicates: the ratio $\underline{G(s)/G(x) > 4}$, main chain scission reactions are predominating, and positive tone pattern with this polymer can be achieved.[176]

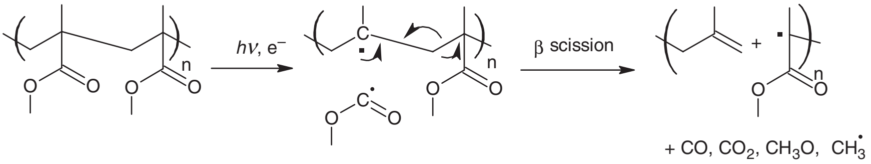
G(*s*) values for positive tone, main chain scission resists can be calculated by plotting the reciprocal of the number average molecular weight of the resist vs absorbed dose. Figure 7.36 provides results for three polymer systems.[177] A resultant linear relationship is obtained and the slope of the line provides for determination of the value of *G*(*s*).

PMMA is one of the most thoroughly studied positive acting electron beam resists, and its radiation induced reaction pathway leading to a reduction in molecular weight is shown in Fig. 7.37.[178]

Main chain scission events occurring in PMMA are initiated via e-beam or short wavelength UV (<250 nm) exposure through cleavage of the main chain carbon to the carbonyl–carbon bond. The resulting radical species,



7.36 Plot of reciprocal number average molecular weight vs absorbed dose for three candidate positive acting e-beam resist materials. The equation of the line requires the knowledge of the initial number average molecular weight M_n^0 of the polymer and the absorbed dose (D) for each irradiation event. The $G(s)$ value for PMMA is provided as a reference point.



7.37 Radiation induced reaction process in polymethylmethacrylate resist.

which are formed along the main chain, undergoes a rapid rearrangement that results in the cleavage of the polymer backbone and production of small molecule volatile products and additional radical species.

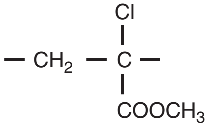
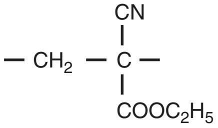
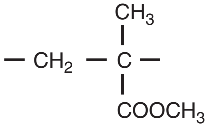
Positive tone pattern formation in PMMA is realized from the dissolution rate of the exposed film regions in a chosen developer exceeding the dissolution rate of the non-exposed film areas. This kinetic rate difference in a given solvent (developer) is the cornerstone contrast mechanism for methacrylate-based e-beam resists.[179] In molecular terms, the dissolution rate of the fragmented polymer chains in the exposed regions is faster than that of the unobstructed polymer chains in the non-exposed material. Since there is a distribution of polymer chain lengths in both the non-exposed and exposed regions, there is likewise a range of dissolution rates as a function of chain length. This process is inherently characterized as being of low to moderate contrast. Contrast can be improved if the molecular weight distribution $\left(\bar{M}_w/\bar{M}_n\right)$ or polydispersity of the starting polymer material approaches unity.[180] A unity polydispersity results in a more uniform dissolution rate of polymer when subjected to electron beam or ionizing radiation.

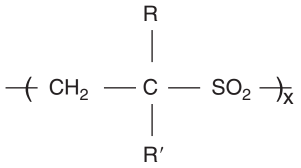
The reported e-beam sensitivity and contrast value of PMMA[179,181] can vary considerably depending on the:

- single- or multi-component solvent used as the developer and rinse
- development time and temperature/RH set points
- mode of development, such as immersion quiescent, immersion with agitation, spray-spin, ultrasonic, or a megasonic process

PMMA has been shown to be a high resolution resist capable of <20 nm resolution when using a very thin (<50 nm) film, high accelerating potential beam energy (>50 keV), and a development process using high frequency (1 MHz) agitation.[182] However, as a consequence of having a relatively low $G(s)$ value of 1.3, PMMA is an insensitive resist, and has only marginal dry

Table 7.4 Examples of two positive acting e-beam resists containing high electronegativity/electron affinity substituents present on the backbone of the polymer. The incorporation of these substituents into the acrylic based resists is shown to improve the $G(s)$ value when compared to PMMA.

Resist	Chemical structure of repeat unit	$G(s)$
Poly(methyl-α-chloroacrylate)	 $-CH_2-\overset{Cl}{\underset{COOCH_3}{C}}-$	5
Poly(ethyl-α-cyanoacrylate)	 $-CH_2-\overset{CN}{\underset{COOC_2H_5}{C}}-$	2
Poly(methyl methacrylate)	 $-CH_2-\overset{CH_3}{\underset{COOCH_3}{C}}-$	1.3



$$-\!\left(\mathrm{CH_2}-\overset{\displaystyle R}{\underset{\displaystyle R'}{\overset{|}{\underset{|}{\mathrm{C}}}}}-\mathrm{SO_2}\right)_x$$

PBS: R = H, R′ = CH_2CH_3
PMPS: R = CH_3, R′ = $CH_2CH_2CH_3$

PBS resist properties
Sensitivity: 0.5–1.0 μC/cm^2@10 keV
Contrast: 1–3
Resolution: 0.50 μm

7.38 Generalized chemical structure of the repeat unit for the 1-olefin sulfone copolymers and the e-beam lithographic performance characteristics for PBS resist.

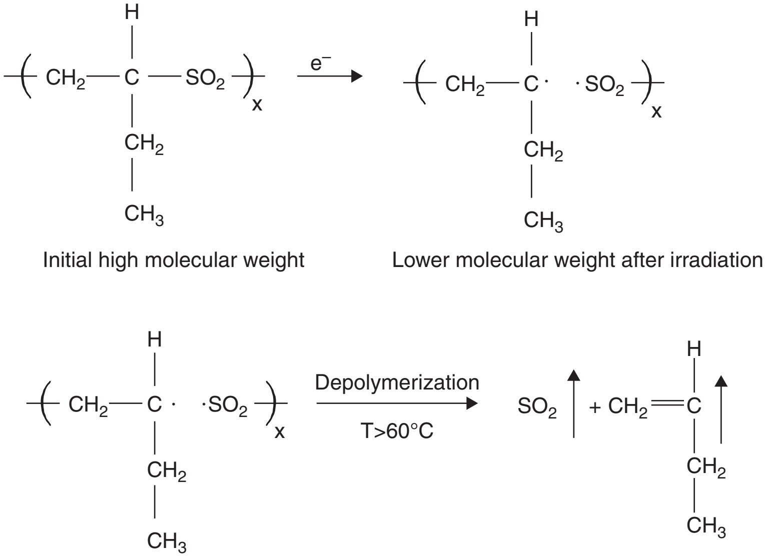
etching resistance in a Cl_2–O_2 plasma, rendering it not useful for commercial optical mask fabrication.[183] Instead, PMMA has found wide utility in direct write on semiconductor substrate applications where resolution, not sensitivity, is the major driver in the choice of the resist.

Structural modifications to PMMA through introduction of polar electronegative substituents/high electron affinity at the quaternary carbon of the backbone or on the ester side chain can improve the main chain scission yield for this class of resists.[184] Table 7.4 lists examples of resists having the substituent present on the backbone of the polymer, and compares the $G(s)$ values to that of PMMA.

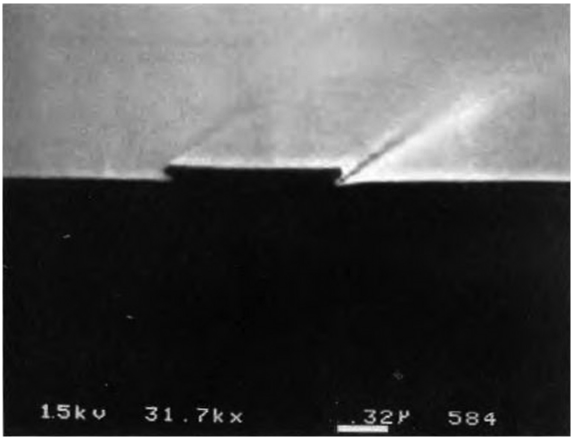
The positive acting main chain scission type resist, which has been shown to exhibit one of the highest electron beam sensitivities, is based on the alternating copolymer of 1-butene and sulfur dioxide (PBS).[185] PBS found wide usage in conjunction with the first commercially available raster scan e-beam exposure tools used in the photomask mask making industry.[186,187,188] These exposure tools required a resist to have a sensitivity/process dose of $<1.0\ \mu C/cm^2$ at 10 keV to enable high throughput operation.[189]

Figure 7.38 provides the generalized chemical structure and lithographic performance for this family of resists. The electron beam induced reaction mechanism in PBS is shown in Fig. 7.39. Mechanistically, electron beam exposure induces the highly efficient homolytic cleavage of the weak main chain carbon–sulfur bond yielding $G(s) > 10$ and high electron beam sensitivity values.[190] The resultant high sensitivity of PBS had historically made it the workhorse positive acting optical mask making resist. However, development process inflexibility and complete lack of plasma dry etching resistance have substantially reduced its usefulness in fabricating advanced lithography technologies masks.[191]

Specifically, in the situation where the dimensions of the CDs printed on the wafer have dropped below 50 nm, it was no longer possible to fabricate

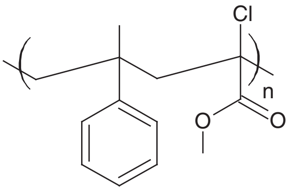

e⁻
Initial high molecular weight
Lower molecular weight after irradiation
Depolymerization
T>60°C

7.39 Electron beam induced degradation mechanism in PBS. The depolymerization process occurs in PBS at temperatures greater than its 60°C T_c.

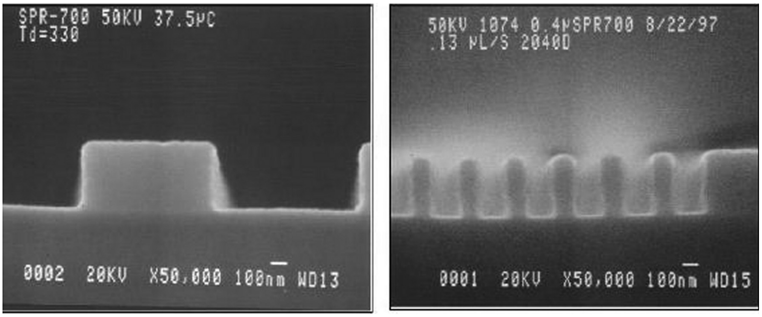

1.5kv
31.7kx
.32P
584

7.40 SEM micrograph of the undercut profile of a < 1.0 coded μm Cr line formed using a wet etch pattern transfer process.

the corresponding photomask using a biased isotropic wet chemical etching pattern transfer process for the chrome layer of the photomask.[192] Figure 7.40 reveals the undercut profile of < 1.0 μm chrome-on-glass pattern formed using a wet chemical etch pattern transfer process. A > 10%



7.41 Chemical structure of the repeat unit in the ZEP family of positive acting electron beam resist.

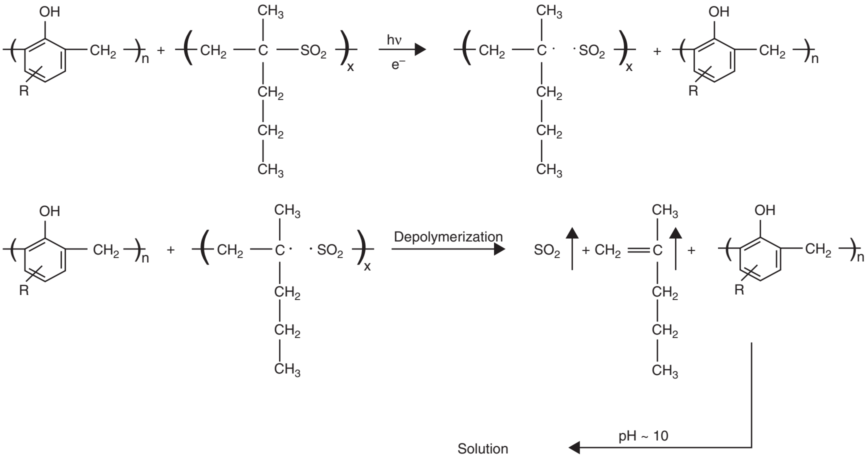

SPR-700 50KV 37.5μC
Td=330
0002 20KV X50,000 100nm WD13
50KV 1074 0.4μSPR700 8/22/97
.13 μL/S 2040D
0001 20KV X50,000 100nm WD15

0.7 μm l/s at 37.5 μC/cm^2 0.13 μm l/s at 40 μC/cm^2

7.42 SEM micrographs of submicron vertical pattern profiles in SPR 700 i-line resist exposed using an electron beam lithography tool operating @ 50 keV.

CD loss is realized from the undercut, and replacement of the wet to dry Cr etch pattern process became necessary. This required the search for an electron beam sensitive resist that provided durability in an O_2–Cl_2 plasma used to dry etch the chrome mask layer. Positive acting resists containing aromatic substituents represented attractive candidates for achieving reduced O_2–Cl_2 plasma etch erosion rates and vertical Cr pattern profiles.

In this regard, a preferred alternative to PMMA and PBS resists is the ZEP family of commercial electron beam resists.[193,194] ZEP affords higher e-beam sensitivity over PMMA and greater dry etch resistance than both PMMA and PBS. The commercially available ZEP family of resists is an organic developable resist based on an alternating copolymer of α-methylstyrene and α-chloromethacrylate. Figure 7.41 provides the chemical structure of the repeat unit in ZEP resist. The α-methylstyrene monomer is incorporated to enhance dry etch durability and the α-chloromethacrylate promotes sensitivity to electron beam radiation. Using ZEP-7000, a dose of 8 μC/cm^2 @ 10 keV was required for patterning ≤1.0 μm Cr mask features. An O_2–Cl_2 plasma etch rate selectivity (etch rate Cr/etch rate ZEP 7000) was 2.9:1.[195]

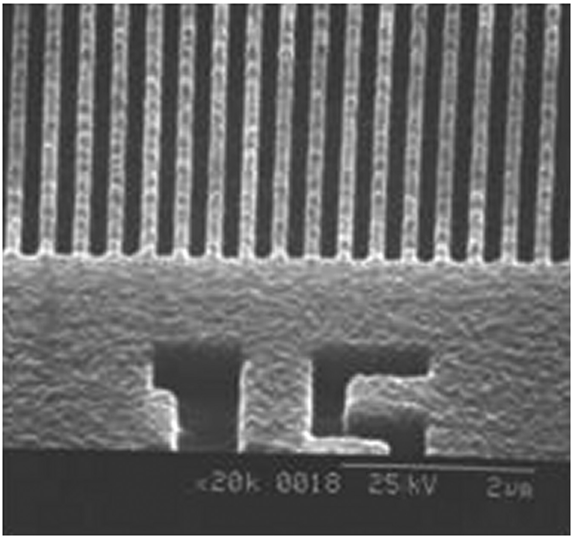


7.43 E-beam induced reaction mechanism in poly (2-methyl-1-pentene-co-sulfur dioxide) PMPS, which has been blended in with a novolac resin.

Commercial versions of novolac-DNQ resists used in i, and g-line lithographies exhibit moderate electron beam sensitivity (20–40 μC/cm^2 at 50keV), O_2–Cl_2 plasma etch rate selectivity > 3, and use a non-organic based developer. Figure 7.42 depicts submicron imaging and vertical profiles in SPR 700 i-line resist exposed at 50 keV.[196]

Enhancement in the e-beam sensitivity of aqueous base developable novolac resists was reported by workers at Bell Laboratories, who investigated a polymer blend composed of 80–95 wt.% m,p cresol novolac and 5-20 wt.% of the 1-olefin sulfone copolymer poly (2-methyl-1-pentene-co-sulfur dioxide) PMPS.[186] In addition, workers at IBM reported blending a novolac with a terpolymer consisting of sulfur dioxide, 2-methyl-1-pentene, and a methallyl ether derivative.[197] The sulfone co- and ter-polymers act as an aqueous base DIs of the novolac resin. In the case of PMPS, which has a low T_c (~−40°C), upon exposure to e-beam radiation, undergoes spontaneous depolymerization, rendering the exposed regions soluble in an aqueous base developer medium.[185] The radiation induced reaction mechanism for this copolymer when blended with a novolac resin is shown in Fig. 7.43. This resist exhibits an e-beam sensitivity in the range of 5–10 μC/cm^2 at 50 keV and as shown in Fig. 7.44 has been used to produce the high resolution line and space patterns.[166]

The further thrust by the photomask industry[183] to identify a positive acting, highly electron beam sensitive, dry etching resist has led to the use of styrene-based CA resists commercialized for use in 248 nm lithography. This, however, has necessitated the mask industry to adapt to the added

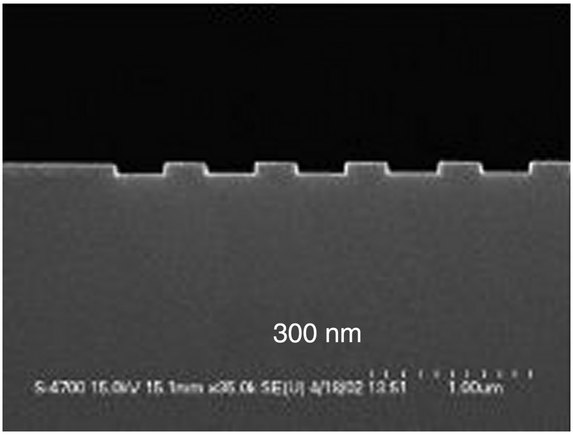


7.44 SEM micrograph of coded 0.15 μm line and space patterns in a PMPS/novolac blended positive acting electron beam resist. Initial resist film thickness = 0.30 μm.

process complexity of baking on thick glass plates and environmental control issues that come with these types of resists.[198] In addition, due to the potential variation in the CA resist lithographic performance as a function of the time the coated mask blank is stored, the mask industry could no longer keep a large inventory of these coated blanks. A number of high- and low-activation energy CA resists for mask making were provided for fabricating <0.50 μm dense lines at doses <5 μC/cm^2 at 50 keV and >1:1 selectivity in the Cr dry etching process. Figure 7.45 provides results in TOK REAP 200 positive acting CA e-beam resist, which supports these performance characteristics.[187]

7.6.3 Positive acting resists for high resolution and throughput electron beam lithography

Electron beam sensitive positive acting resists based on chemical amplification are the platform of choice for projection electron and multi e-beam lithography technologies developed for high throughput and resolution applications.[170,172,199] Resist performance requirements for the ≤80 nm device generations are comparable to what is needed in advanced high wafer throughput optical lithography technologies. Table 7.5 provides requirements of a resist for use in a high throughput electron lithography technology system.



7.45 SEM micrograph of profiles of coded 300 nm line and space patterns in 70 nm thick Cr mask layer using TOK REAP 200 CA resist and a Cl_2-O_2 plasma dry etch pattern transfer process.

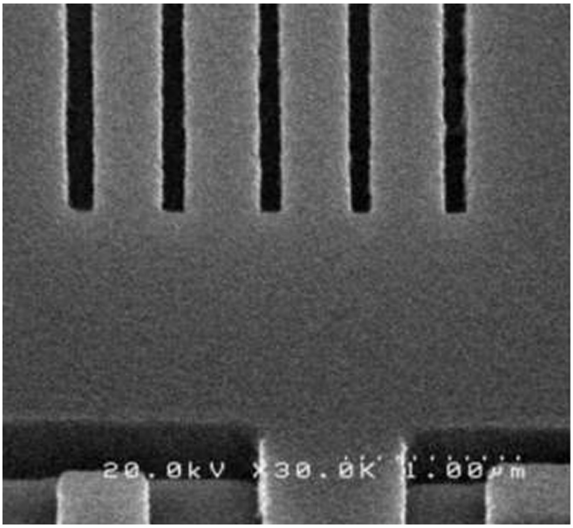
Table 7.5 Property vs performance requirements for resists use in a high throughput electron beam lithography exposure technologies

Property	Requirement
Nominal L/S feature size(nm)	80–35
Process/type	SLR/CA
Image tone	+
Process stability	No outgassing
Aspect ratio/profile (°)	3–4/87–90
Process dose (μC/cm^2)@ 100 keV	5–10
Dose latitude (%range)	>10
Contrast	>5
LER (nm), 3σ	<2
PED stability (h)	>3
PEB sensitivity (nm/°C)	<2
Dev. time sensitivity (nm/min)	<2
Etch selectivity	=PHS[a]

[a] PHS-Polyhydroxystyrene.

Identification of a highly sensitive resist allows for exposures to operate at lower beam currents, which reduces electron–electron interactions and beam blurring. This in turn improves resist resolution and process latitude.[200] Figure 7.46 is an example of <100 nm isolated trench resist patterns imaged at a dose of 5.4 μC/cm^2 at 100 keV using the SCALPEL system.[201]

A low resist process dose can also reduce exposure induced wafer heating and expansion, thereby decreasing overlay errors.[202] The potential for a rise

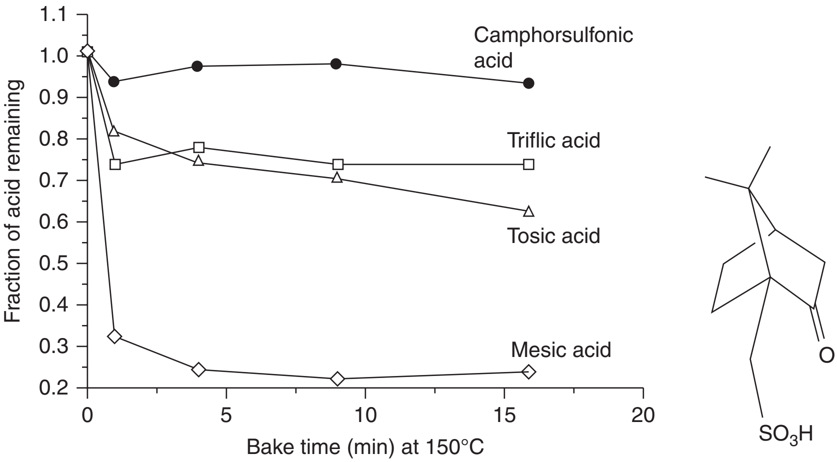


7.46 SEM micrograph of coded 80 nm isolated trench patterns formed in 360 nm thick experimental TOK positive acting CA resist using the electron beam projection lithography system (SCALPEL). Exposure dose = 5.4 μC/cm^2@ 100 keV.

in substrate temperature during the high vacuum exposure environment of a high throughput e-beam lithography system puts a constraint on the type of CA resist that can be used. As in the case of CA resists that are compatible with the operation of optical lithography tools, the optimum e-beam resist formulation is one in which the exposure energy requirement is of a practical minimum, and the exposure solely results in the generation of the acid molecules, which do not volatize at the maximum wafer temperature ($T_{\max}$) realized from the absorption of the incident beam of electrons into the substrate.[203]

As previously described, identifying a bulky/large molecule acid precursor and radiation generated acidic product that can associate with the polar functionalities of the resin component of the resist represents a successful approach to eliminating acid loss. Figure 7.47 provides a measure of the thermal stability for a number of acids investigated for inclusion in CA resist formulations developed to be compatible with the operation of high throughput DWEBL systems.[204]

In addition, to further eliminate volatile product evolution, the subsequent acid catalyzed reaction should ideally have a kinetic rate constant (k_d) ~ 0 at $T_{\max}$. Choice of the CA resist in the case of the acid catalyzing a thermally assisted deprotection reaction should have a high activation energy (E_a). If K_d is not equal to zero, then the volatility of the acid catalyzed reaction



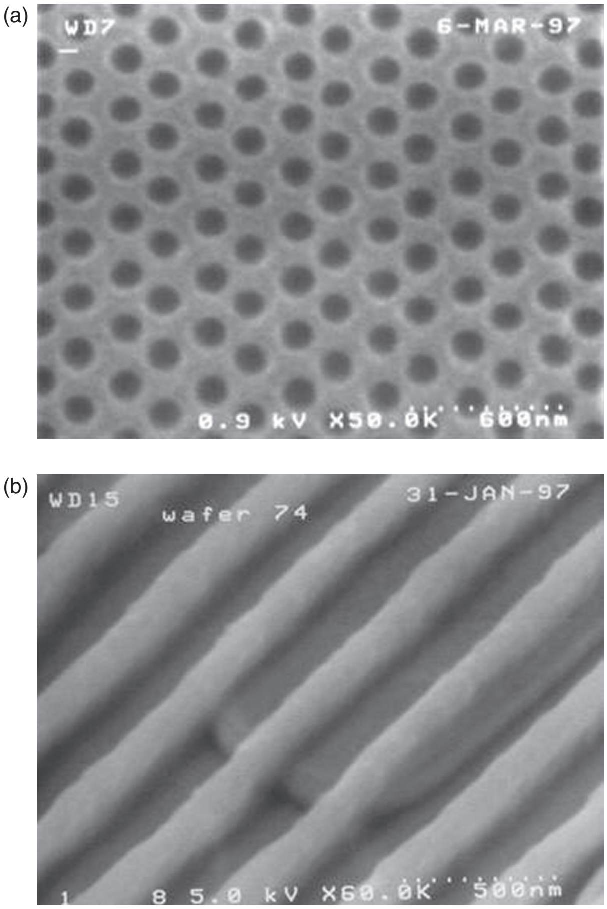
7.47 Fraction of the mass of the acid remaining vs heating (bake) time for a number of acids investigated for use in CA positive acting resist formulations. The chemical structure of camphorsulfonic acid is provided as an example of one acid species which meets the low volatility requirement when present in an e-beam exposure tool.

products must be zero. Choice of the protecting group determines the level of E_a and Fig. 7.15 is a summary of where select protecting groups fall on the E_a scale.

Optimization of the compatibility of the CA resist with the operating conditions of the e-beam lithography tool enables the IC fabrication engineer to exploit the additional advantages of the technology. This includes the ability to use:

- A true single resist layer process, thereby removing the need for an anti-reflexive coating (ARC),
- The large DOF to print <100 nm contact holes in thick resist,[170]
- The printing of resist patterns over existing topography while maintaining the required CD of those patterns.[205]

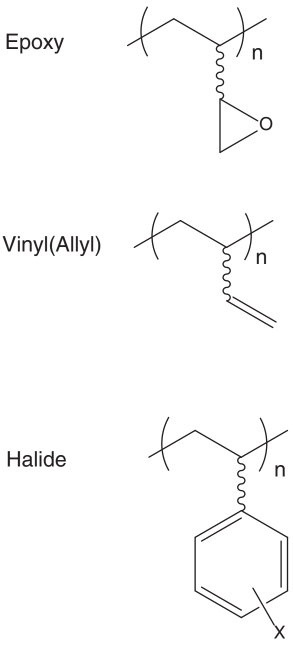
Figure 7.48 provides examples of the last two described advantages. Further refinements in e-beam sensitive CA resists have established this lithography technology as a candidate for use in <22 nm device generation.[188] However, in the case of direct e-beam write patterning of substrates where resolution and process simplicity are key requirements, non-CA resists have also been widely exploited. These include self-assembled monolayer materials, inorganic resists such as metal halides, and oxides, which all have been shown to provide pattern resolution below 10 nm.[9]



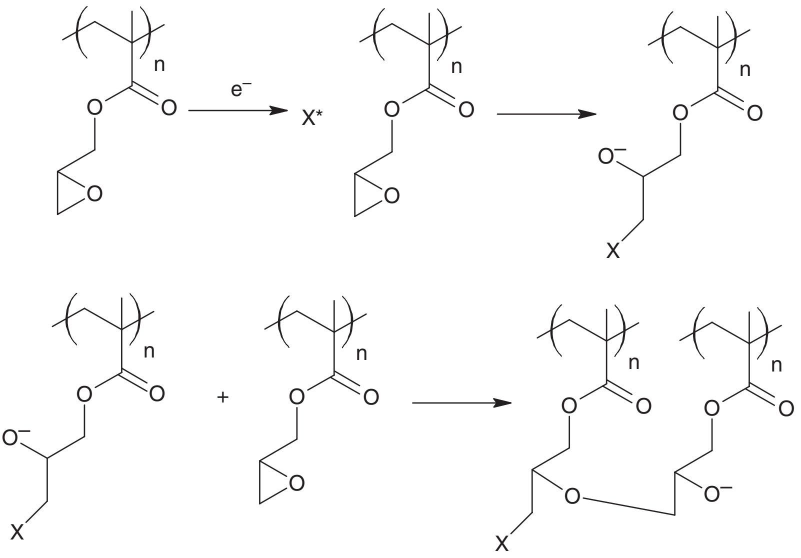
7.48 (a) SEM micrograph depicting a dense array of coded 80 nm contact holes patterned in a 750 nm thick positive acting CA resist and (b) coded 180 nm resist lines and spaces printed over 0.8 μm wide Tungsten Silicide (WSi) lines which are 0.20 μm in height. The measured deviation in resist line width on and off the underlying topography is 0.04 nm. Resist patterned was performed using the SCALPEL system.

7.6.4 Negative acting non-chemically amplified electron beam resists

It has been previously stated that when the ratio of $G(s)/G(x) < 4$, the predominant electron beam radiation induced reaction in the resist is inter-chain crosslinking and leads to negative tone resist formation. Early version electron beam resists used to fabricate 'clear field' optical masks


Epoxy
n
O
Vinyl(Allyl)
n
Halide
n
X

7.49 Examples of side chain/pendent groups on an acrylate backbone polymer which are used as initiation sites for inter-chain crosslinking reactions in negative acting e-beam resists.



7.50 Electron beam induced reaction pathway in PGMA negative acting resist.

and to pattern isolated line structures on silicon substrates were based on acrylate or methacrylate backbone polymers and copolymers.[206,207] These materials contained side chain/pendent groups that acted as sites through which intermolecular crosslinking between polymer chains and intramolecular crosslinking on the same polymer chain could occur. Examples of side chain groups incorporated into negative resists to promote inter-chain crosslinking reactions are shown in Fig. 7.49.

In the case of epoxy- and allyl-containing resists, negative tone formation occurs through the electron beam radiation induced reaction pathway, which initially forms radical and/or cationic species that reacts with like groups on adjacent polymer chains to form an inter-chain bond or cross-link. For example, e-beam exposure of the epoxy-containing homopolymer PolyGlycidylMethacrylate (PGMA) leads to ring opening of the epoxide ring, which then reacts with an epoxy ring of an adjacent polymer molecule.[208] This process continues via a chain reaction leading to the formation of an insoluble polymer network. This represents a form of amplification process whereby one radiation induced event produces a cascade of ring opening, followed by inter-chain bond formation, events.[209] Figure 7.50 details the e-beam induced reaction sequence in PGMA. This reaction process has resulted in epoxy-containing negative resists exhibiting high sensitivity of <1.0 μC/cm^2 at 10 keV.[206]

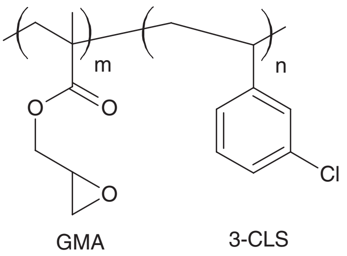
A drawback to this type of resist is that the propagation of the radiation generated reactive species continues in the vacuum environment of the e-beam exposure tool. This results in those patterns exposed at the onset having dimensions (greater) than the patterns exposed last.[210] This can result in difficulties in meeting the required CD control for the device being produced.

Key materials properties of negative acting crosslinking polymeric resists that affect its sensitivity are:

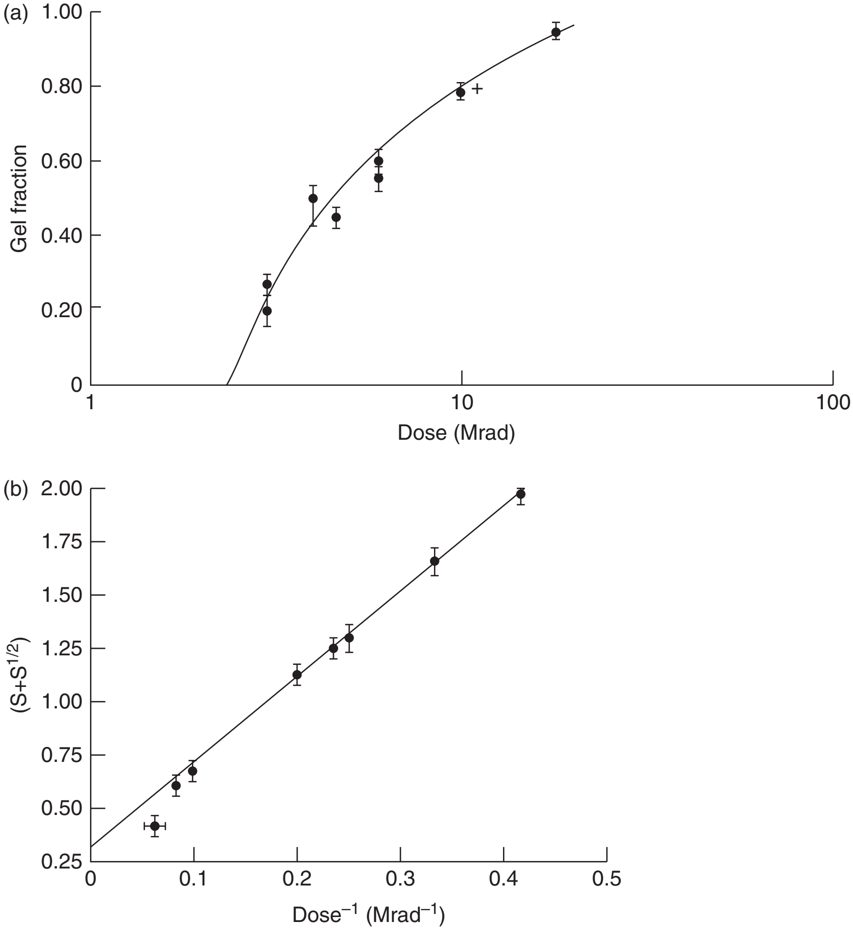
- Weight average molecular weight $\left(\bar{M}_w\right)$
- Molecular weight distribution/polydispersity $\left(\bar{M}_w/\bar{M}_n\right)$
- Composition

In general, the higher the $\bar{M}_w$, the higher is the sensitivity, and the narrower the $\left(\bar{M}_w/\bar{M}_n\right)$ distribution, the higher is the contrast of the resist.[175, 211] For copolymers, varying the composition offers the flexibility of tuning the desired resist properties, such as high sensitivity and high dry etch resistance, into a single material.[210,212,213]

A classical methodology used to determine if a polymer will act as a negative acting resist is to subject powdered samples of the polymer present in a degassed, sealed under vacuum, glass tube to Cobalt-60 γ irradiation followed by an exhaustive solvent soxhlet extraction process.[174,214] The



7.51 Chemical structure of the repeat unit in poly(glycidyl methacrylate-co-3-chlorostyrene) (GMC) negative acting electron beam resist.



7.52 (a) Plot of gel fraction vs absorbed dose for the GMC copolymer having a $\bar{M}_w$ of 1.94×10^5 g/mol and containing 46 mol% of the monomer glycidyl methacrylate (GMA). (b) Charlesby-Pinner Plot for GMC using the data shown in Fig. 7.52*a*.

amount of any insoluble material remaining after the extraction process has been completed is then used to calculate the gel/insoluble fraction remaining from the starting material. Evidence of a gel fraction indicates that the $G(s)/G(x)$ ratio <4 and the radiation induced inter-chain crosslinking reaction predominates over the dose range the material was subjected to.

This methodology was used to investigate the suitability of the copolymer of glycidyl methacrylate and 3-chlorostyrene (GMC) as a sensitive negative acting e-beam resist.[173] The chemical structure of the repeat unit in GMC is provided in Fig. 7.51. The post-gamma-irradiation gel fraction as measured by the soxhlet extraction method vs absorbed dose for the GMC is plotted in Fig. 7.52a.

For polymers such as GMC, which have a random distribution, knowledge of the soluble fraction as a function of absorbed dose allows one to make use of the Charlesby–Pinner relationship[215]

$$S + S^{1/2} = \frac{G(s)}{2G(x)} + \left(\frac{9.65\times 10^5}{\bar{M}_w}\right)\left(\frac{1}{G(x)}\right)\left(\frac{l}{r}\right)$$

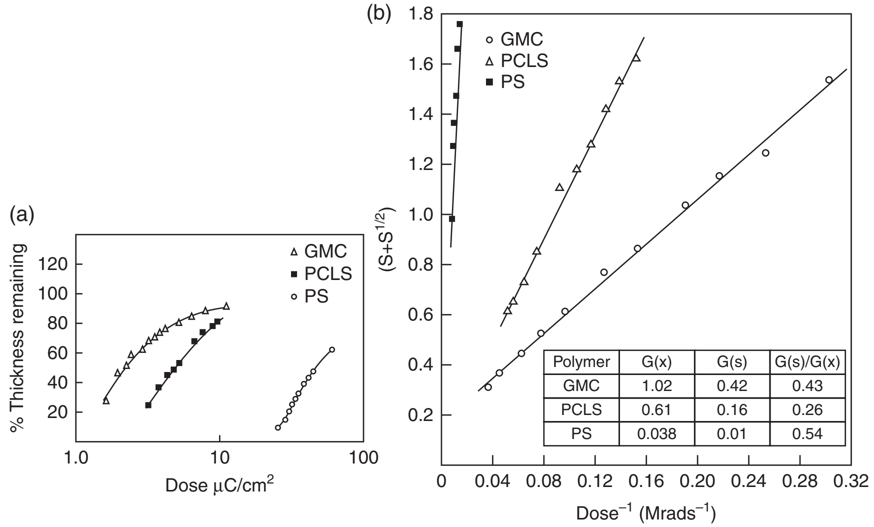
where S = soluble fraction, $G(s)$ = G-value scission, $G(\mathrm{x})$ = G-value crosslinking, $\bar{M}_w$= weight average molecular weight (initial) in grams/mole, and r = absorbed dose in Mrads.

A resulting linear relationship between the soluble fraction term ($\mathrm{S} + \mathrm{S}^{1/2}$) and the reciprocal absorbed dose (l/r) is used to calculate the ratio of $G(s)/G(x)$, and the individual $G(s)$ and $G(x)$ values. Figure 7.52b makes use of the data in the plot above it to construct the Charlesby–Pinner plot for GMC.

From this plot, values of 1.13 and 0.65 are calculated for $G(x)$ and $G(s)$, respectively. The ratio of $G(s)/G(x) = 0.58$ which is <<4, enabling one to conclude GMC as a sensitive negative acting e-beam resist. Alternatively, the data from electron beam exposure response curves of known negative acting resists can be used to produce a subsequent Charlesby–Pinner plot. This method[173,213] is useful as a relative comparison of the calculated G-values, and in the determination of which radiation induced reaction pathway $G(x)$ or $G(s)$ predominates over the dose range applicable for lithography patterning purposes.

Figure 7.53 shows the electron beam exposure response curves and corresponding Charlesby–Pinner plots for three negative acting resists; GMC, poly 3-chlorostyrene (PCLS), and polystyrene (PS). The resultant radiation chemical yields for the three negative acting resists are given in the table embedded in Fig. 7.53b.

From this relative comparison, $G(s)/G(x)$ is lowest for the chlorinated aromatic polymer PCLS. The e-beam induced radiation degradation mechanism for a chloromethylated aromatic polymer (PCMS) is outlined in



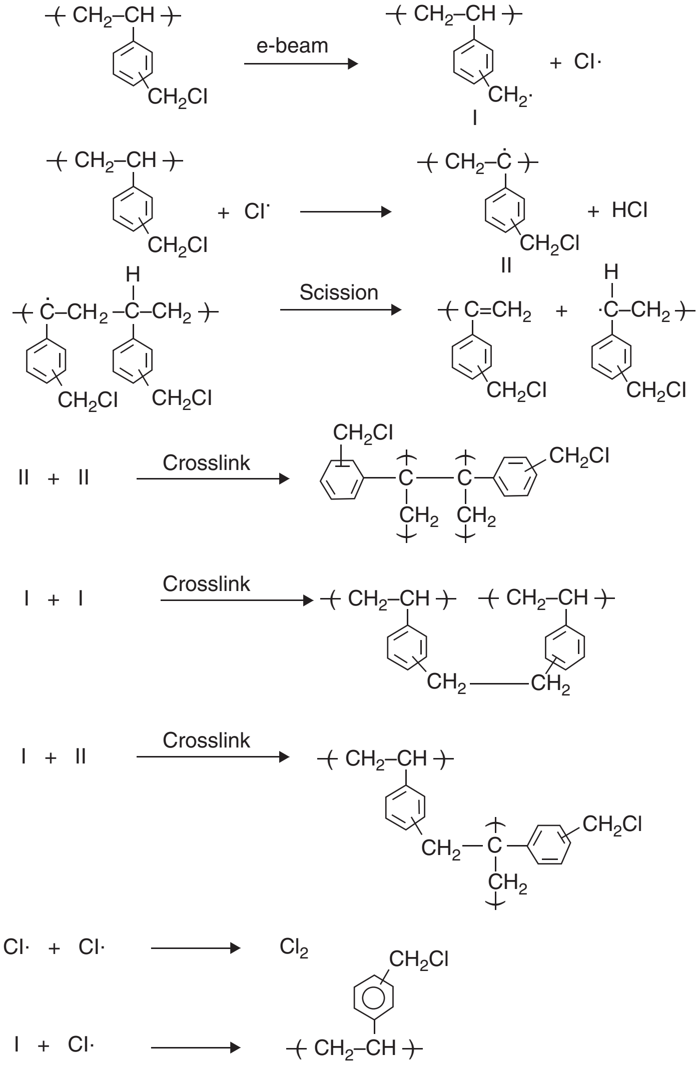
Polymer	G(x)	G(s)	G(s)/G(x)
GMC	1.02	0.42	0.43
PCLS	0.61	0.16	0.26
PS	0.038	0.01	0.54

7.53 (a) E-beam exposure response curves for GMC, poly-3-chlorostyrene (PCLS) and polystyrene (PS) resists. (b) Corresponding Charlesby-Pinner plots and *G* values using data from Fig. 7.53a.

Fig. 7.54. The reaction pathway is characterized by the initially formed radicals combining to become directly involved in the subsequent inter-chain crosslink reactions, thereby terminating any further propagation of chemical events.[216] As shown in Fig. 7.54, the number of inter-chain crosslinking events substantially exceeds the singular reaction event resulting in main chain scission. This supports the relatively low calculated $G(s)/G(x)$ ratio for halogenated styrene based negative resists and higher contrast over vinyl- or epoxy-containing methacrylate-based resists.[217–219] The lack of any propagation of the radiation induced chemical events by the initially formed radicals also eliminates any variability in the size of the resist image while the exposed substrate remains in the vacuum environment of the e-beam lithography tool.[220]

Table 7.6 lists a number of chlorinated styrene resists and select material and lithographic properties.[184] The table also shows the strong dependence of resist sensitivity on the $\bar{M}_w$ of the material. Further enhancement in the performance of PCMS has been realized by preparing material with extremely narrow polydispersity.[221] Figure 7.55 shows the inverse relationship between the contrast value of PCMS and the polydispersity of the polymer. This suggests that narrowing the polydispersity could improve the resolution capability of the resist. This figure also suggests that starting with lower molecular weight material improves contrast; however, this comes at the expense of resist sensitivity. Narrow MWD PCMS resist has found wide utility in the production of clear-field optical masks as well in direct write on Si wafer e-beam lithography applications.[183]

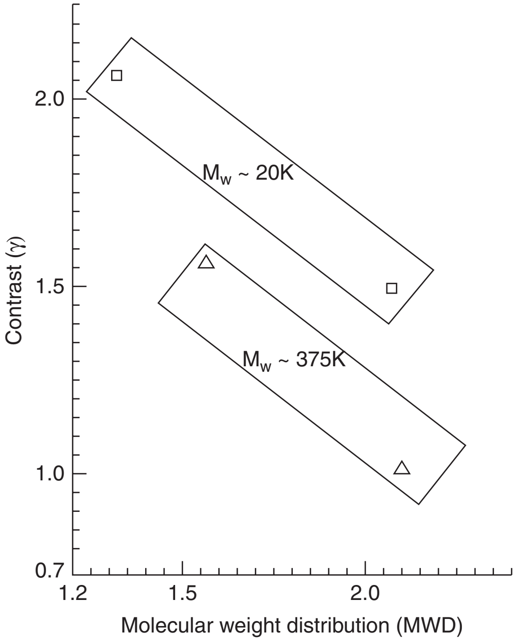
G(*s*) values have also been used to determine the stability of a resist in an oxygen (O_2) plasma etching environment.[222] Table 7.7 provides the relative O_2 plasma etch rate of known positive acting electron beam resists to that of polystyrene. The data support the statement: the higher the value



7.54 Electron beam induced reaction sequence occurring in the chlorinated aromatic negative acting resist: Poly m,p-chloromethyl styrene (PCMS).

Table 7.6 Examples of chlorinated styrene resists, their molecular properties and corresponding electron beam sensitivity

Resist	Mw	E-beam sensitivity @ 20kV
Poly(chloromethylstyrene) (PCMS)	$2 \times 10^4 - 4.5 \times 10^5$	7–0.35
Chloromethylated Poly (α -Methyl Styrene) (~95% Chloromethylated)	$8 \times 10^3 - 1.9 \times 10^5$	42–3
Poly(4-Chlorostyrene) (PCS)	$3 \times 10^5 - 7 \times 10^5$	2.5–1.5

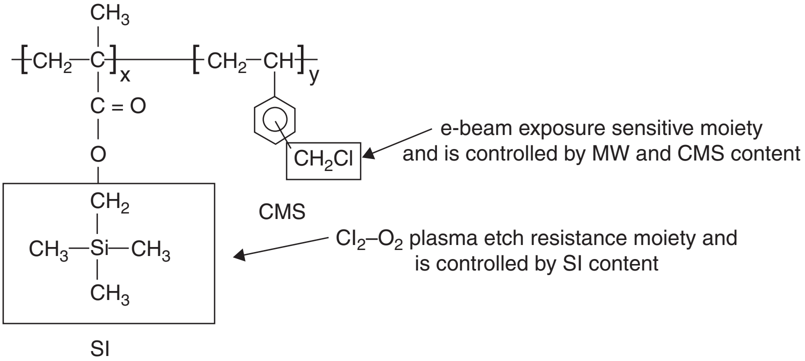


7.55 Plot of the contrast value for the e-beam sensitive PCMS resist vs the molecular properties of the starting polymer material.

of $G(s)$, the higher is the O_2 plasma etch rate of the polymer. The correlation of plasma etch resistance to the $G(s)$ value has been shown not to be a function of the type of main chain scission pathway, but dependent only on the magnitude of $G(s)$. This early study led to the pursuit of identifying e-beam sensitive resists having low $G(s)$ and $G(s)/G(x)$ values, and to contain aromatic moieties to minimize film erosion during the dry etching of the Cr layers of the photomask.[167,195,223,224] A follow-on generation of resists used in mask making further improved the O_2–Cl_2 dry etching resistance, through the incorporation of silicon into the structure of the

Table 7.7 Oxygen plasma dry etch rate ratio (etch rate resist/etch rate polystyrene) and *G*(*s*) values for PS and select positive acting electron beam resists

Polymer	O_2 Plasma dry etch rate ratio	*G(s)*
Poly(styrene)	1.0	0.01
Poly(methyl methacrylate)	2.4	1.2
Poly(methyl methacrylate-co-methacrylonitrile) 94:6	2.7	2.0
Poly(1-butene sulfone)	7.1	9.0


CH3
CH2
C
x
CH2
CH
y
C = O
O
CH2
CH3—Si—CH3
CH3
SI
CH2Cl
CMS
e-beam exposure sensitive moiety and is controlled by MW and CMS content
Cl2–O2 plasma etch resistance moiety and is controlled by SI content

7.56 The chemical structure of the repeat unit in P(SI-CMS) and description of the lithographic properties imparted by the highlighted functional groups present in the copolymer.

material.[223,225] A representative resist is shown in Fig. 7.56, and is based on a copolymer of m,p-chloromethylstyrene and trimethylsilylmethyl methacrylate P(SI-CMS).[226] This negative acting resist continues the usage of the chloromethylstyrene monomer for its promotion of electron beam sensitivity, and the silicon methacrylate monomer for its improvement in dry etching resistance in an O_2-Cl_2 plasma. P(SI-CMS) formulations having the molecular, compositional and thermal properties summarized in Table 7.8 were used to fabricate optical masks. High sensitivity and vertical side wall photomask patterns could be achieved when using a P(SI-CMS) formulation having an $\bar{M}_w > 1.0 \times 10^5$ g/mol and containing 90 mol% trimethylsilylmethyl methacrylate (SI).[225]

A Cr dry etch pattern transfer process using P(SI-CMS) resist and the corresponding etch selectivity values, as compared to novolac-based resists, are provided in Tables 7.9 and 7.10, respectively. A greater than five-fold improvement in etch selectivity is realized through the use of the organosilicon resist vs a standard novolac material. Fig. 7.57 displays submicron

Table 7.8 Material properties of a range P(SI-CMS) formulations used in the fabrication of Cr-on-glass optical masks

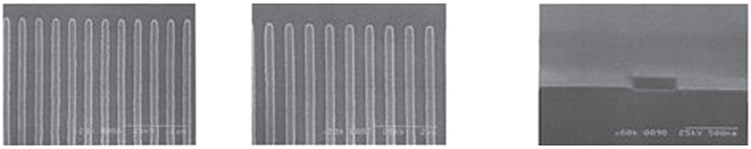
Material property	Value for P(SI-CMS)
Weight average molecular weight	$0.30–1.3 \times 10^5$ g/mol
Molecular weight distribution	1.4–1.6
Composition	90 mol% SI
Glass transition temperature (T_g)	82–90°C
Thermal decomposition temp. (T_d)	>200°C

Table 7.9 Sequence of steps and conditions used in the P(SI-CMS) Cr dry etch optical mask pattern transfer process

Process step	Conditions
Post-develop bake	80–110°C/30 min
Descum	Fluorine based plasma
Oxidation	Oxygen plasma
Cr etch	Chlorine-Oxygen plasma

Table 7.10 Etch rate and selectivity values of P(SI-CMS), novolac vs the optical mask Chrome layer

Material	Etching rate (Å/min)	Etching rate ratio (Cr rate/P(SI-CMS) resist rate)
$P(SI_{0.90}–CMS_{0.10})$	20	4.25
Novolac	280	0.30
Chrome	85	1.0



0.25 μm Cr lines and spaces

0.30 μm Cr lines and spaces

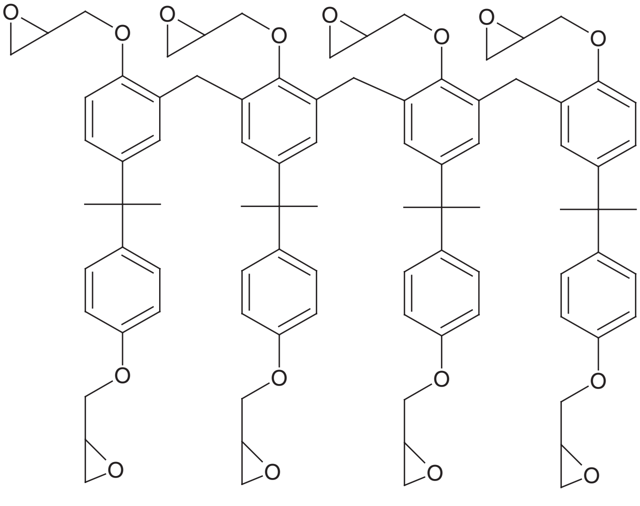
Profile of a isolated 0.5 μm Cr line

7.57 SEM micrographs of coded ≤ 0.50 μm line and space and isolated line Cr patterns.

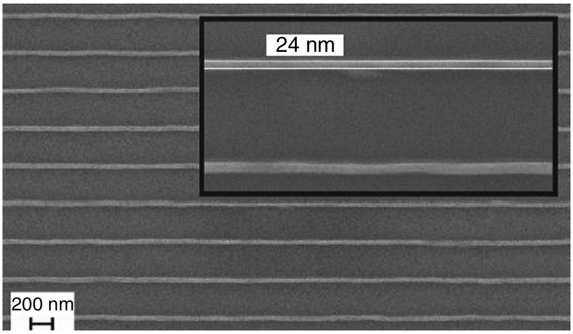
vertical Cr line profiles formed when using this resist, making this a suitable material for fabricating ≤250 nm device generation optical masks.

7.6.5 Negative acting chemically amplified electron beam resists

Consistent with non-CA electron beam resists, the predominant radiation induced reaction mechanism to yield a negative image in CA resists is based on crosslinking. As described earlier, these CA resists undergo crosslinking via an acid catalyzed condensation reaction between a PHS based polymer and a melamine crosslinking agent or a ring opening reaction followed by a series of subsequent reactions with the same ring moiety on an adjacent oligomeric or polymeric chain. The latter negative acting CA type resist was initially developed for thick (≥200 μm) layer applications used in MEMS and LIGA x-ray lithography processes.[227,228] The resist referred to as SU-8 is based on a multi-functional epoxy novolac resin[229] and commercialized versions of this two-component resist have been made available by MicroChem Corp. A representative chemical structure of the resin component of the resist is provided in Fig. 7.58. SU-8 has also been investigated for use in the fabrication of optical masks[230] and for high resolution electron beam lithography processes. Exposure of this resist initially forms a Brønsted acid, which protonates an epoxide group present in the oligomer. The protonated oxonium sites react during the PEB step with



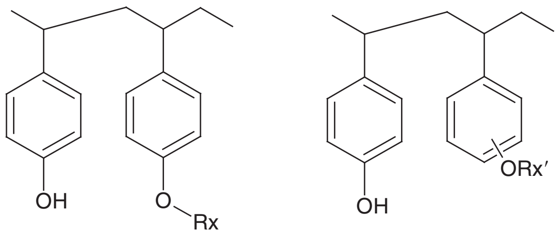
7.58 Chemical structure of the oligomeric resin component in SU-8 resist. An acid generating species is added to the resin to effect the formation of negative tone resist images.



7.59 SEM micrograph of sub-25 nm line patterns formed in a 100 nm thick SU-8 film when exposed to 100 keV electron beam radiation.

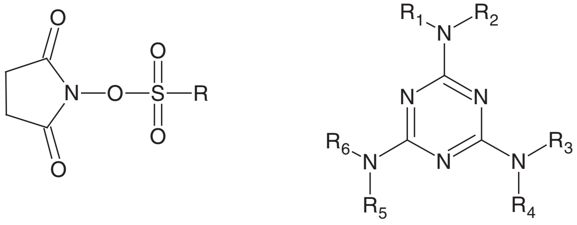
epoxide groups present on different oligomer chains to produce a highly crosslinked network. This renders the exposed regions of the film insoluble in the organic based developer used to delineate the resist pattern.[231] The resist can be further 'hardened' or more extensively crosslinked by subjecting the patterned film to a high temperature (>150°C) post-development bake. Once hardened, the resist exhibits high thermal stability and exceptional dry etching resistance. Electron beam sensitivity in the 10–15 μC/cm^2 range at 100 keV, and sub-25 nm images in 100 nm thick SU-8 as shown in Fig. 7.59 have been reported when used in direct write substrate patterning applications.[232]

Multi-component, aqueous base developable, negative acting electron beam sensitive CA resists containing an acid generating species, a multi-functional melamine crosslinking agent, and a PHS based polymer or blend of resins of this type were utilized in various e-beam lithography approaches.[233,234] Figure 7.60 provides representative chemical structures for the components present in this resist, which, when exposed to e-beam radiation, generates an acid and initiates a condensation reaction between the PHS resin and the melamine crosslinking agent. These materials were commercially available from Shipley (SAL resist platform) and Sumitomo Chemical (NEB resist platform). Table 7.11 summarizes the lithographic performance of the NEB version 22A resist when exposed to 50 and 100 keV electrons of a DWEBL or a PEL system.[235,236] The range in sensitivity of this resist was a function of the process baking conditions and high resolution, high aspect ratio patterns[236] are shown in the Fig. 7.61. Both SU-8 and SAL/NEB negative CA resists provide for appreciable halogen-based plasma etch resistance over that observed with acrylate-based resists. These resists have been used as etch masks in deep Si trench patterning processes.[237,238]



Polymer A

Polymer B



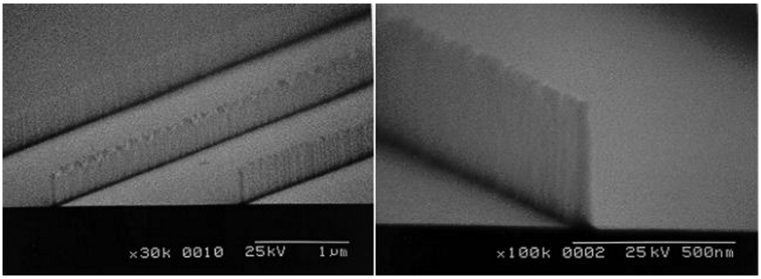
Photoacid generator

Crosslinker

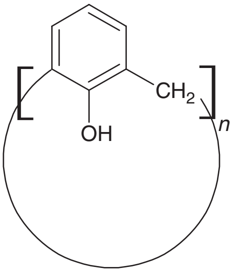
7.60 Chemical structure of the components in a negative acting electron beam sensitive resist which can be developed in aqueous alkaline solution.

Table 7.11 Lithographic property vs performance results for Sumitomo NEB 22A negative acting resist when exposed to 100 keV electrons

Lithographic property	Performance
Film loss after exposure	~0
Stability in exposure tool vacuum environment	>24 h
Sensitivity	15–30 μC/cm^2
Contrast	5–7
Resolution	0.10 μm isolated lines in 0.40 μm thick resist (PEL), 0.03 μm in 0.10 μm thick resist (DWEBL)



7.61 SEM micrographs of high aspect ratio (>6) sub-100 nm isolated lines imaged in NEB 22A resist when exposed using a 50 keV electron beam lithography system.



7.62 Generalized chemical structure of calixarene. Molecules having $n = 4,6$ and 8 units have been most widely investigated for use as a high resolution negative acting electron beam resist.

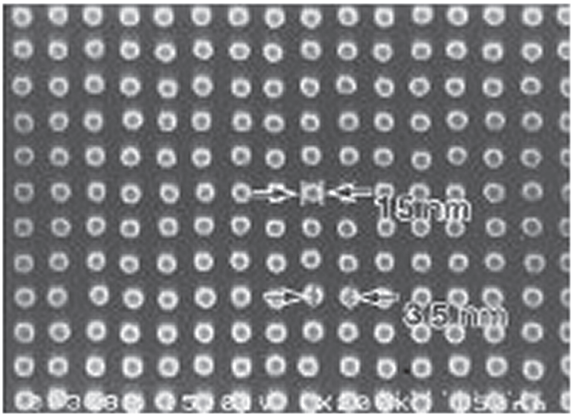
The following sections describe non-CA negative acting electron beam resists such as calixarene and Hydrogen SilsesQuioxane (HSQ), which have been extensively investigated for use in direct write-on-silicon substrate applications when resolution capability is the overriding requirement.

7.6.6 Calixarenes

Calixarenes belong to a group of organic macrocyclic molecules involving the linking up a series of phenolic units through methylene bridges at the ortho-position to the hydroxyl group present on the aromatic moiety.[239] Figure 7.62 provides the generalized chemical structure of calix[n]arene, where in general the number of phenolic units ranges from 4–20. Calix[n]arenes are characterized as having high thermal stability ($T_m > 300°C$) and do not degrade in a variety of chemical environments. Calix[n]arenes, however, possess limited solubility in organic solvents and only through derivatization of the parent molecule do improvements in solubility occur. An example of this is the work reported for (hexaacetate p-methylcalix[6]arene).[240]

Calix[n]arenes have a small molecular size compared to polymeric-based resists, which affords the ability to produce line patterns below 10 nm and dot arrays on the order of 15 nm. Figure 7.63 displays a 15 nm size dot array in a 30 nm thick calixarene resist.[241] However, electron beam dose requirements to produce images at this length scale are in the mC/cm^2 at 50 keV range.

Modifications to the structure to enhance the sensitivity to electron beam radiation have been shown for the calixarene derivative, hexachloromethyl-hexacetoxycalix[6]arene. A dose of 0.7 mC/cm^2 at 50 keV was used for patterning of <50 nm sized features.[242] Simplification in the processing of these materials has been realized using calixarene p-chloromethyl methoxy-calix[4]arene, available from Tokuyama Corp under the commercial name



7.63 SEM micrograph of a coded 15 nm dot array in a calixarene resist patterned using a direct write e-beam lithography system.

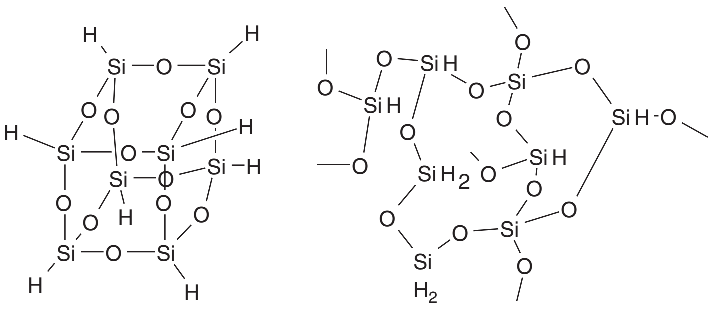
TEBN-1. This calixarene acts as a thermally developable negative acting resist over a 250–400°C temperature range.[243] The lack of the use of an organic solvent developer improves the overall resist image quality resulting in the delineation of highly dense sub-30 nm half-pitch line and space patterns.

7.6.7 Hydrogen silsesquioxane

HSQ is an inorganic oligomeric material belonging to the family of spherosiloxanes. It is characterized as having a low dielectric constant (~2.8), and spin coated films can act as a planarizing interlayer dielectric film layer of an IC device.[244] HSQ possesses good mechanical stability, and the relatively compact nature of the molecule affords the ability to fabricate ~10 nm line pattern features having smooth edge profiles when exposed in a DWEBLS.

HSQ has been made commercially available by Dow Corning and is promoted as a flow on oxide. Solutions of HSQ in methyl isobutyl ketone, when stored at 5°C, are expected to have a shelf life of ~6 months. Maintaining solutions at higher temperatures can increase the $\bar{M}_w$ and MWD of the dissolved HSQ, resulting in its e-beam lithographic performance being tied to solution storage conditions.[245]

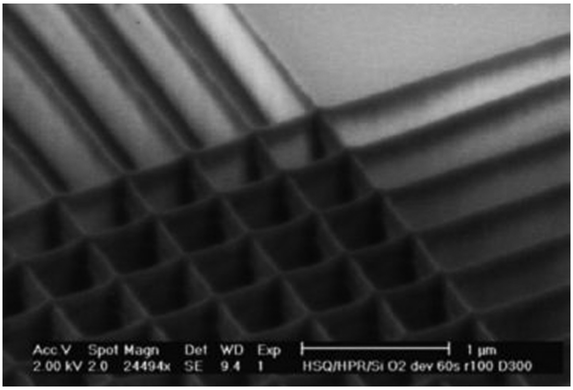
Namastu *et al.*[246] have postulated an e-beam induced radiation degradation pathway for HSQ, whereby the initial event involves cleavage of a Si–H bond in the material. At this cleavage site, radical species reacts with water present in the spin coated film to produce highly reactive silanol groups. The silanol groups undergo condensation reactions to



Cage form

Network form

7.64 Chemical structure for the cage and network forms of HSQ inorganic resist.


Acc V Spot Magn Det WD Exp
2.00 kV 2.0 24494x SE 9.4 1
1 µm
HSQ/HPR/Si O2 dev 60s r100 D300

7.65 SEM micrograph of orthogonal coded 50 nm equal lines and spaces formed in a bilayer resist stack comprised of a top 140 nm thick HSQ and a lower 660 nm thick hard bake novolac film.

form Si–O–Si linkages, which convert HSQ from a cage structure into a network form. Representative chemical structures of the cage and network forms of HSQ are shown in Fig. 7.64. An alternative HSQ crosslinking mechanism has been reported[247] using data collected from obtaining vibrational and electron beam induced desorption spectra of the irradiated material.

Ultra-high resolution (≤30 nm) patterning in thin single layer HSQ resists can be achieved using moderate to high keV electron beam radiation and development in a 0.26 N TMAH solution.[248]

The high silicon content of HSQ allows it to be highly resistant in an oxygen plasma. The combination of high resolution and dry etch resistance has enabled the patterning as shown in Fig. 7.65 of sub-50 nm wide, high aspect ratio features when using a bilayer resist scheme.[249]

7.6.8 Hafnium and zirconium based resists

The replacement of SiO_x with HfO_x as the gate dielectric layer of an IC device, and the goal to simplify the patterning process for this material has led to the development of a spin coatable resist containing HfO_x and $HfSO_x$.[7,250] The patterning of the HfO_x layer can be simplified by removing the etching step through direct exposure of that spun on resist layer. A sol-gel based spin coatable HfO_x resist derived from reacting hafnium *tert*-butoxide with benzoyl acetone has been shown to be sensitive to optical and electron beam radiation.[251] Exposure leads to a reduction in the overall organic component of the exposed regions of the resist film, and development in ethanol leads to negative pattern formation. An electron beam dose of 115 mC/cm^2 at 100 keV is required to insolubilize 50% of a 130 nm thick film. Use of higher exposure doses has resulted in the patterning of ~10 nm wide lines having very low line edge roughness (LER).

Inorganic resists based on hafnium and zirconium oxide sulfates ($HfOS_x$, $ZrOS_x$) improve the e-beam sensitivity of this class of materials. This can be achieved without scarifying the ability to pattern nanoscale size features and development in an aqueous alkaline solution.

Insolubilization of the exposed film areas is a result of the decomposition and elimination of peroxide, which promotes crosslinking through metal–oxo bridges and the formation of 3D framework structures. Sub-30 nm wide features have been formed in a $ZrOS_x$ negative acting resist using a 50 μC/cm^2 at 30 keV exposure dose and aqueous base developer.[252] These metal oxides resists also exhibit CHF_3 plasma etch rates which are a factor >7 times lower than thermally grown oxide.

7.7 Resists for selected forward looking lithographic technologies

To push resolution to a smaller scale, new lithography technologies have been proposed for 16 nm (and beyond) device generations. Resist materials in these forward looking lithographic technologies can significantly deviate from the CA resists used in DUV lithography, and many of them are being developed for use beyond the electronics industry. Three lithographic technologies of particular interest are nanoimprint, dip-pen/polymer pen, and directed self-assembly. The last technology is the topic of discussion in Chapter 8, and readers are referred to that chapter for a comprehensive discussion on this forward looking technology.

7.7.1 Nanoimprint lithography

Thermal nanoimprint lithography (T-NIL)

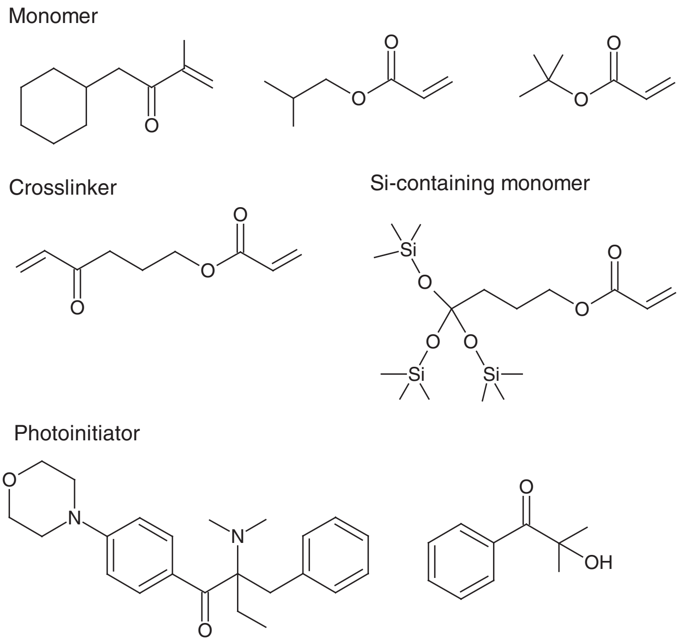
In the T-NIL (see Section 9.4 of Chapter 9) process, a relief pattern is generated by applying pressure to a rigid patterned mold/stamp located in contact with the top of a thermal plastic polymer film heated to a temperature above the T_g of the polymer. The mold replicated relief pattern initially present in the rubbery state of the polymer is preserved when the polymer film is cooled to room temperature, and the stamp is released from the film.

T-NIL resist materials need to have a T_g below the temperature used during the molding process, and mechanical strength/high Young's modulus to withstand the room temperature (RT) mold release process. Commercial versions of varying MW (25–950 k g/mol) PMMA having a $T_g \geq 100°C$ were used as the resist material in the T-NIL process. PMMA enables modest molding temperatures to be used, and exhibits acceptable dry etch resistance during the final step of the pattern formation process.[253] Polystyrene, polycarbonates, and cyclo-olefin copolymers exhibit desirable mechanical and dry etch resolution properties, and have also been explored as the resist medium in the T-NIL process.

UV nanoimprint lithography (UV-NIL)

In UV-NIL (see Sections 9.6–9.7 of Chapter 9), a UV light transparent patterned stamp presses and passes light into selected areas of a photo-curable liquid resist, which undergoes a crosslinking reaction in the material. After the exposure induced crosslinking reaction is completed, the mold is released and the mechanical structure of the mold is preserved in the resist. The RT resist curing process eliminates the lengthy thermal cycle experienced in T-NIL. The initial liquid state of the UV curable resist also prevents thermal distortion and mismatch of thermal expansion coefficients between mold and substrate (which can occur in T-NIL), thereby enabling good overlay and alignment.

An ideal UV-NIL material should exhibit low viscosity for optimizing the material dispensing and spreading process, high UV photo-crosslinking efficiency, ease of release from the stamp, and appreciable etch resistance after it is crosslinked. Typical UV-NIL resist formulations consist of a bulk photo-polymerizable monomer, such as an acrylate or vinyl ether, a silane/siloxane-containing monomer to promote oxygen etch resistance, and a crosslinking agent to provide mechanical strength and thermal stability of the imprinted structure. Figure 7.66 illustrates a few examples and corresponding chemical structures of the components of UV-NIL resist formulations.[254–256]



7.66 Chemical structures of select monomers, photoinitiator and crosslinking species incorporated into UV-NIL resists.

7.7.2 Dip-pen/polymer pen nanolithography

Dip-pen nanolithography (DPN) is a scanning probe-based molecular level printing lithography technology[257] that applies an atomic force microscopy (AFM) tip to directly form both soft and hard (<50 nm) resist patterns.[258] DPN deviates from conventional photolithography in that it does not require an exposure tool and a mask, and has been developed as a unique and complementary lithographic tool for the fabrication of biological based devices. Early demonstrations of DPN consisted of depositing 15 nm wide features of an alkylthiol such as 1-octadecanethiol (ODT) onto an Au substrate. DPN technology has evolved to generate high resolution patterns in soft materials, such as DNA and proteins, small organic molecules, and polymers, and in hard materials, such as sol-gels, semiconductors, and metal oxide nanoparticles.[259]

DPN, however, can suffer from an intrinsic throughput limitation typical for AFM tip based delivery and deposition technologies. To overcome the throughput issue, a cantilever based multi-million-array nanolithography technology was introduced to print a massively parallel array in two-dimensions.[260] Later, a cantilever-free scanning probe molecular printing technology, referred to as polymer pen nanolithography (PPL), was introduced to overcome the throughput limitation of DPN and the use of complicated pen arrays.[261] Through a combination of microcontact printing and a scanning probe approach, a soft poly(dimethylsiloxane) (PDMS)-based elastomeric pyramid is used as a multi-million-array lithography tip to deliver and deposit resist materials on a substrate. Materials such as alkylthiols[262] and proteins[263] have been reported for use in the PPL technology.

7.8 Resist resolution limitations

Upon closer review of Fig. 7.1, the IC manufacturer is seen to be well along the way of introducing into production IC devices having minimum CDs of ≤22 nm. Concurrent with this shrinking of the CD is the need for the resist to exhibit high exposure sensitivity (Table 7.2) and performance reproducibility across 450 mm diameter Si substrates. There are, however, limitations to what can be achieved with commercial CA resist formulations that currently must be used to meet the IC manufacturer's product throughput requirements.

Three main factors affecting how far these resists can be used are resist LER, resist pattern collapse, and resist blur. LER is a main topic covered in Chapter 16, whereas the latter two factors are discussed in the following sections.

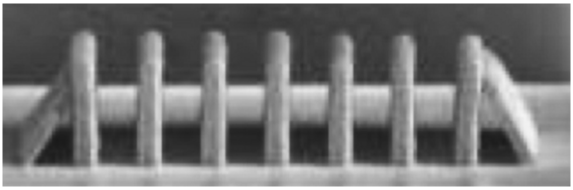
7.8.1 Resist pattern collapse

An issue preventing high resolution resist pattern images is the collapsing of adjacent features after the wet process development and drying steps.[264] Pattern collapse occurs when unbalanced capillary forces generate from outside to inside the pattern during the drying step of the resist. Factors influencing the resist pattern to collapse or not, include the Young's modulus of the resist material, stress in the spun on film, and adhesion of the resist-coated film onto the substrate.

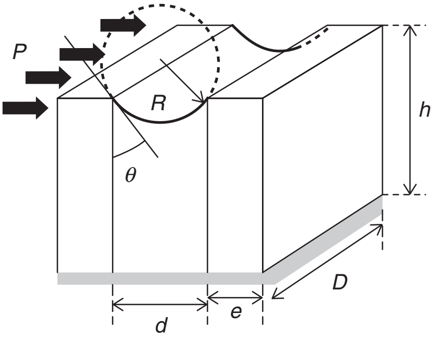
The resist image collapse process is pictorially described in Fig. 7.67[265] and the following equations are used to derive the pressure difference within the resist (P) and the force on the resist pattern (F):

$$P = \frac{2\gamma\cos\theta}{d}$$

$$F = PDh = \frac{2Dh\gamma\cos\theta}{d}$$



7.67 SEM cross sectional image of a resist pattern after completion of the wet development and drying steps. Note the collapsing of the terminal line patterns onto the adjacent interior lines.



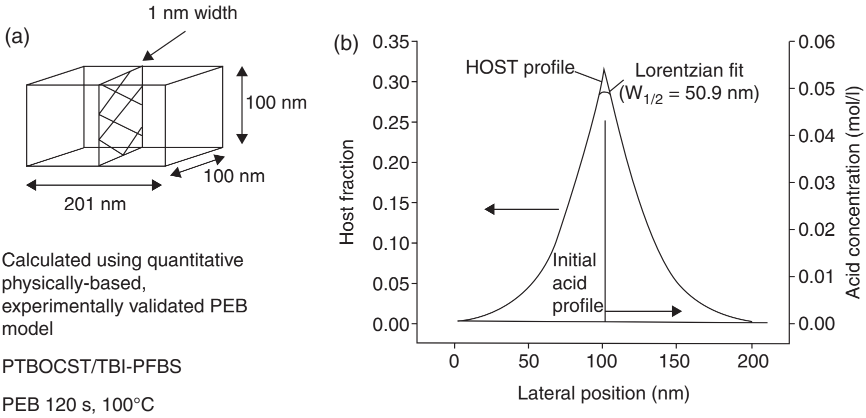
7.68 Diagrammatic representation of the resist pattern collapse process.

where D is the length of the resist line, h is the resist height, d is the spacing distance between adjacent resist lines, e as depicted in Fig. 7.68 is the resist feature width, γ is the surface tension, and θ is the meniscus angle along resist wall.

In other words, during development, capillary forces form inside the confined areas between adjacent line features as a consequence of the surface tension of the meniscus of the liquid present within that area. Pattern collapse occurs when horizontal forces overcome the mechanical forces of resist. Practical solutions for reducing resist pattern collapse include adding surfactant into the rinse liquid[266] and reducing the aspect ratio (h/e) of the resist line to 2 or less.

7.8.2 Resist blur

During the photo exposure process, the aerial image transfers into the resist as a latent image of the photo generated acid. This in turn creates an acid gradient between the exposed and unexposed areas of the resist film. In the subsequent PEB step, the kinetics of the diffusion of the acid molecules originating from such a gradient competes with that



7.69 (a) Illustrates a model for the aerial image blurring processes in a CA resist for the case when a 1 nm wide line of uniformly exposed latent acid image embeds between adjacent unexposed areas. (b) Shows a comparison of the initial resist aerial image to a broadened acid distribution profile resulting from the diffusion of that species during the process PEB step.

of the acid deprotection reaction of the polymer. Once the acid at the boundary of exposed and unexposed area diffuses, the deprotection reaction continues to an undesirable extent, thereby deteriorating the latent image contrast and increasing the observance of the blurring of the resist image. Figure 7.69a illustrates a model for these processes in a CA resist when compared with the initial acid profile; Figure 7.69b shows a broadened acid distribution that is caused by acid diffusion upon PEB process. The resulting developable latent acid profile can be modeled as a truncated Lorentzian function with a full width at half maximum height of 50.9 nm.[267] A further modeling effort to compare the effect of acid diffusion and deprotection on overall resist blur has shown that acid diffusion becomes the dominant component in resist blur when the line width shrinks to 50 nm or less. At this dimension, additional steps must be taken in the design of resists which exhibit 'less blur.'[267]

7.9 Conclusion

This chapter has provided the fundamental lithography technology node-specific approaches to designing and synthesizing resist materials and their associated patterning processes. These approaches have been challenged by the requirement of addressing the concomitant demands of shrinking the critical dimension of the IC device and increasing the density of those elements in the device.

Current thrusts in developing double patterning and double exposure resist processes have pushed the envelope of what can be done with the most advanced 193 nm lithography tools. Continued refinement of EUVL and development of non-radiation-based lithography technologies offer options to meeting the next generation nanolithography requirements. However, future disruptive changes to the makeup of the resist medium may be required to sustain the ever-vanishing CD size and tolerance ranges demanded by the IC manufacturers.

7.10 References

1. Chiang C and Kawa J (2007), *Design for Manufacturability and Yield for Nano-Scale CMOS*, Berlin, Springer.
2. Sivakumar S (2006), 'Lithography challenges for 32 nm technologies and beyond', *Proc IEDM,* San Francisco, California, December 10–13, 2006, 1–4.
3. Thompson L F, Willson C G and Bowden M J (1983), *Introduction to Microlithography*, Washington D.C., ACS Symposium Series 219, 1–14, American Chemical Society.
4. Reichmanis E and Novembre A E (1993), 'Lithographic resist materials chemistry', *Annual Review of Materials Science*, **23**, 11–43.
5. Krysak M, De Silva A, Sha J, Lee J- K and Ober C K (2009), 'Molecular glass resists for next generation lithography', *Proc SPIE,* **7273**, 72732N-1–8.
6. Lawson R A, Tolbert L M and Henderson C L (2010), 'Single-component molecular resists containing bound photoacid generator functionality', *J Micro/Nanolitho MEMS MOEMS*, **9**, 013015.
7. Trikeriotis M, Bae W J, Schwartz E, Krysak M, Lafferty N, Xie P, Smith B, Zimmerman P A, Ober C K and Giannelis E P (2010), 'Development of an inorganic photoresist for DUV, EUV, and electron beam imaging', *Proc SPIE,* **7639**, 76390E-I-10.
8. Bowden M J (1984), *Materials for Microlithography*, Washington, D.C., ACS Symposium Series 219, 39–117, American Chemical Society.
9. Grigorescu A E and Hagen C W (2009), 'Resists for sub-20-nm electron beam lithography with a focus on HSQ: state of the art', *Nanotechnology,* **20**, 292001.
10. Novembre A E, Ocola L E, Houlihan F, Knurek C and Blakey M (1998), 'New developments in resist materials for the SCALPEL technology', *J Photopolym Sci Tech,* **11**, 541–546.
11. Kunz R R (2004), 'Photoresist outgassing: a potential achilles heel for short-wavelength optical lithography', *Proc SPIE,* **5376**, 1–15.
12. Helbert J N (2001), *Handbook of VLSI Microlithography: Principles, Technology, and Applications,* Norwich, NY, Noyces Publications/Williams Andrew Publishing, LLC, p. 698.
13. Gokan H, Esho S and Ohnishi Y (1983), 'Dry etch resistance of organic materials', *J Electrochem Soc,* **130**, 143–146.
14. Mogab C J (1983), '*Plasma Etching*', *VLSI Technology*, 1st ed. M. Sze, ed., Singapore, McGraw Hill, 303–345.
15. Moritz H (1985), 'Optical single layer lift-off process', *IEEE Trans Electron Devices,* **32**, 672–676.

16. Madou M J (2001), *Fundamentals of Microfabrication: The Science of Miniaturization,* 2nd ed. Boca Raton FL, CRC Press, 20.
17. Reiser A (1989), *Photoreactive Polymers: The Science and Technology of Resists,* New York, John Wiley and Sons.
18. Mack C (2007), *Fundamental Principles of Optical Lithography: The Science of Microfabrication*, England, John Wiley and Sons.
19. Nonogaki S, Ueno T and Ito T (1998), *Microlithography Fundamentals in Semiconductor Devices and Fabrication Technology*, New York, Marcel Dekker, Inc.
20. Sheats J R and Smith B W (1998), *Microlithography Science and Technology*, New York, Marcel Dekker, Inc., 429–514.
21. Kishimura S, Yamaguchi, A, Yamada Y and Nagata H (1992), 'Masking effect of photoactive compounds with various ballast molecules in novolak-naphthoquinonediazide positive photoresists', *Polym Eng Sci*, **32**, 1550–1555.
22. Furuta A, Hanabata M and Uemura Y (1986), 'High performance positive photoresists', *J Vac Sci Technol B*, **4**, 430–436.
23. Süs O (1944), 'Über die Natur der Belichtungsprodukte von Diazoverbindungen. Übergänge von aromatischen 6-Ringen in 5-Ringe', *Justus Liebigs Annalen der Chemie* **556**, 65–84.
24. Pacansky J (1980), 'Recent advances in the photodecomposition mechanism of diazo-oxides', *Polym Eng Sci*, **20**, 1049.
25. Suzuki K and Smith B W (2007), *Microlithography: Science and Technology*, 2nd ed., CRC Press, Taylor and Francis Group, Boca Raton, FL, 602.
26. Pol V, Bennewitz J H, Escher G C, Feldman M, Firtion V A, Jewell T E, Wilcomb B E and Clemens J T (1986), 'Excimer laser-based lithography: a deep ultraviolet wafer stepper', *Proc SPIE,* **633**, 1–16.
27. Dammel R (1993), *Diazonapthoquinone-Based Resists,* Bellingham, WA, SPIE, 179.
28. Ito H and Willson C G (1984), *Polymers in Electronics,* T. Davidson ed., Washington, D.C., ACS Symposium Series 242, 11–23, American Chemical Society.
29. McKean D, Schaedeli U and MacDonald S A (1989), 'Acid photogeneration from sulfonium salts in solid polymer matrices', *J Polym Sci: Part A: Polym Chem*, **27**, 3927–3935.
30. Lamola A A, Szmanda C R and Thackeray J W (1991), 'Chemically amplified resists', *Sol State Tech*, **34**, 53.
31. Reichmanis E, Houlihan F M, Nalamasu O and Neenan T X (1991), 'Chemical amplification mechanism for microlithography', *Chem Mater,* **3**, 394.
32. MacDonald S A, Willson C G and Frechet J M J (1994), 'Chemical amplification in high-resolution imaging system', *Acct Chem Res,* **27**, 151.
33. Crivello J V (1983), 'Possibility for photoimaging using onium salts', *Polym Eng Sci,* **23**, 953.
34. Nalamasu, Reichmanis E, Cheng M, Pol V, Kometani J M, Houlihan F M, Neenan T X, Bohrer M P, Mixon D A, Thompson L F and Takemoto C H (1991), 'Preliminary lithographic characteristics of an all-organic chemically amplified resist formulation for single layer deep-UV lithography', *SPIE,* **1466** 13–25.
35. Maltabes J, Holmes S J, Morrow J R, Barr R L, Hakey M and Reynolds G (1990), '1x deep-UV lithography with chemical amplification for 1-mircron DRAM production', *Proc SPIE,* **1262**, 2–7.

36. Pawlowski, Dammel R, Lindley C R, Merrem H -J, Röschert H and Lingnau J (1990), 'Chemically amplified DUV photoresists using a new class of photoacid-generating compounds', *Proc SPIE,* **1262**, 16–25.
37. Croffie E, Yuan L, Cheng M, Neureuther A, Cirelli R, Watson P, Nalamasu O and Gabor A (2000), 'Modeling influence of structural changes in photoacid generators on 193 nm single layer resist imaging', *J Vac Sci Technol B,* **18**, 3340–3344.
38. Ito H (2000), 'Chemical amplification resists: history and development within IBM', *IBM J Res Develop,* **44**, 119–130.
39. Fouiassier J P and Rabek J F eds, (1993), *Radiation Curing in Polymer Science and Technology* (*Vol II*)*: Photoinitiating Systems,* New York, Elsevier, Ch.8.
40. Cronin M F, Adams T, Fedynyshyn T, Georger J, Michael Mori J, Sinta R and Thackeray J W (1994), 'Investigation of onium salt type photoacid generators in positive DUV resist systems', *Proc. SPIE,* **2195**, 214–224.
41. Allen D A, Opitz J, Larson C E, Wallow T I, DiPietro R A, Breyta G, Sooriyakumaran R, and Hofer D C (1998), *Progress in 193-nm Single Layer Resists: The Role of Photoacid Generator Structure on the Performance of Positive Resists,* Washington D.C., ACS Symposium Series, American Chemical Society.
42. Lamanna W M, Kessel C R, Savu P M, Cheburkov Y, Brinduse S, Kestner T A, Lillquist G J, Parent M J, Moorhouse K S, Zhang Y, Birznieks G, Kruger T and Pallazzotto M C (2002), 'New ionic photo-acid generators (PAGs) incorporating novel perfluorinated anions', *Proc SPIE,* **4690**, 817–831.
43. Tagawa S, Nagahara S, Iwamoto T, Wakita M, Kozawa T, Yamamoto Y, Werst D and Trifunac A D (2000), 'Radiation and photochemistry of onium salt acid generators in chemically amplified resists', *Proc SPIE,* **3999**, 204–214.
44. Crivello J V (1989), 'Novel negative working photoresist', *J Electrochem Soc,* **136**, 1453–1456.
45. Houlihan F M, Shugard A, Gooden R and Reichmanis E (1988), 'Nitrobenzyl ester chemistry for polymer processes involving chemical amplification', *Macromolecules,* **21**, 2001–2006.
46. Neenan T X, Houlihan F M, Reichmanis E, Kometani J M, Bachman B J and Thompson L F (1990), 'Photo- and thermochemistry of select 2,6 dinitrobenzyl esters in polymer matrixes: studies pertaining to chemical amplification and imaging', *Macromolecules,* **23**, 145.
47. Hayashi N, Schlegel L, Ueno T, Shiraishi H and Iwayanagi T (1991), 'Polyvinylphenols protected with tetrahydropyranyl group in chemical amplification positive deep-UV resist systems', *Proc SPIE,* **1466**, 377.
48. Yamaoka T, Nishiki M, Koseki K and Koshiba M (1989), 'A novel positive resists for deep UV lithography', *Polymer Eng Sci*, **29**, 856–858.
49. Schlegel T, Ueno H, Shiraishi N, Hayashi N and Iwayanagi T (1990), 'Acid formation and deprotection reaction by novel sulfonates in a chemical amplification positive photoresist', *Chem Mat,* **2**, 299–305.
50. Onishi Y, Niki H, Kobayashi Y, Hayase R and Oyasato N (1991), 'Acid catalyzed resists for KrF excimer laser lithography', *J Photopolym Sci Technol,* **4**, 337.
51. Shirai M and Tsunooka M (1990), 'Photochemistry of imino sulfonate compounds and their application to chemically amplified resists', *J Photopolym Sci Technol,* **3**, 301.
52. Novembre A E, Hanson J E, Kometani J M, Tai W W, Reichmanis E, Thompson L F and West R J (1992), 'Arylmethyl sulfones: a new class of photoacid generators', *Polymer Eng Sci,* **32**, 1476–1480.

53. Crivello J V (1984), *Polymers in Electronics,* T. Davidson ed., Washington, D.C., ACS Symposium Series 242, 3–10, American Chemical Society.
54. Stewart K, Hatzakis M, Shaw J, Seeger M and Neummann E (1989), 'Simple negative resist for deep ultraviolet, electron beam, and x-ray lithography', *J Vac Sci Technol B,* **7**, 1734.
55. Dubois J C, Eranian A and Datmanti E (1978), 'New photoresists containing thiirane groups', *J Electrochem Soc,* **125**, C154.
56. Hatzakis M, Stewart K J, Shaw J M and Rishton S A (1991), 'New high-resolution and high sensitivity deep UV, x-ray and electron beam resists', *J Electrochem Soc,* **138**, 1076–1079.
57. Allen R D, Conley W and Gerlome J D (1992), 'High-speed aqueous-developing negative resist based on triflic-acid-catalyzed epoxy polymerization', *Proc SPIE,* **1672**, 513–525.
58. Conley W, Moreau W M, Perreault S, Spinillo G T, Wood R L, Gelorme J D and Martino R M (1990), 'Negative tone aqueous developable resist for photon, electron, and x-ray lithography', *Proc SPIE,* **1262**, 49.
59. Feely W E, Imhof J C and Stein C M (1986), 'The role of the latent image in a new dual image, aqueous developable, thermally stable photoresist', *Polym Eng Sci,* **26**, 1101–1104.
60. Thackeray J W, Orsula G W, Pavelchek E K, Canistro D, Bogan L E, Berry A K and Graziano K A (1989), 'Deep UV ANR photoresists for 248 nm excimer laser photolithography', *Proc SPIE,* **1086**, 34–46.
61. Lingnau J, Dammel R and Theiss J (1989), 'Recent trends in X-ray resists: Part II', *Solid State Technol,* **32**, 107–111.
62. Liu H-Y, de Grandpre M P and Feely W E (1988), 'Characterization of a high-resolution novolak-based electron-beam resist with 4 μC/cm^2 sensitivity', *J Vac Sci Technol B,* **6**, 379.
63. Thackerary J W, Orsula G W, Rajaratnam M M, Sinta R, Herr D and Pavelchek E (1991), 'Dissolution inhibition mechanism of ANR photoresists crosslinking vs. -OH site consumption', *Proc SPIE,* **1466**, 39–52.
64. Fréchet J M J, Matuszczak S, Reck B, Stöver H D H and Wilison C G (1989), '*Nonswelling Negative Resists Incorporating Chemical Amplification*', ACS Symposium Series 412, 74–85, American Chemical Society.
65. Reck B, Allen R E, Tweig R J, Willson C G and Matuszczak S (1989), 'Novel photoresist design based on electrophilic aromatic substitution', *Polym Eng Sci,* **29**, 960–964.
66. Fahey J T and Fréchet J M J (1991), 'A new aqueous base-developable negative-tone photoresist based on furans', *Proc SPIE,* **1466** 67–74.
67. Schaedeli U, Holzwarth H, Muenzel N and Schulz R (1992), 'Aqueous base developable deep–UV resist based on chemically amplified crosslinking of phenolic resin', *Polym Eng Sci,* **32**, 1523–1529.
68. Sekiguchi A, Miyake Y and Isono M (2000), 'Analysis of deprotection reaction in chemically amplified resists using an fourier transform infrared spectrometer with an exposure tool', *Jpn J Appl Phys,* **39**, 1392–1398.
69. Yoshino H, Yamana M, Takimoto M and Tanabe H (1999), 'High contrast chemically amplified 193 nm resist for gigabit dynamic random access memory generation', *J Vac Sci Technol B,* **17**, 3322–3325.
70. Seeger D (1997), 'Chemically amplified resists for advanced lithography; road to success or detour', *Solid State Technol,* **40**, 115–118.

71. Hesp S A M, Hayashi N and Ueno T (1991), 'Tetrahydropyranyl- and furanyl-protected polyhydroxystyrene in chemical amplification systems', *J Appl Polym Sci,* **42**, 877–883.
72. Jiang Y and Bassett D R (1992), 'Chemically amplified deep UV photoresists based on acetal-protected poly (vinylphenol)', *ACS Polym Mater Sci Eng,* Preprint, **66**, 41.
73. Ito H, Breyta G, Hofer D and Sooriyakumaran R (1994), 'Environmentally stable chemical amplification positive resist: principle, chemistry, contamination resistance, and lithographic sensitivity', *J Photopolymer Sci Technol,* **7**, 433–448.
74. Dössel K F, Huber H L and Oertel H (1986), 'Highly-sensitive novolak-based x-ray resist', *Microelectron Eng,* **5**, 97–104.
75. O'Brien M J (1989), 'Novolac-based photoresists combining chemical amplification and dissolution inhibition', *Polym Eng Sci,* **29**, 846–849.
76. Novembre A E, Tai W W, Kometani J M, Hanson J E, Nalamasu O, Taylor G N, Reichmanis E and Thompson L F (1992), 'Radiation-induced chemistry of poly(4-[(tert-butoxycarbonyl)oxy]styrene-co-sulfur dioxide)', *Chem Mater,* **4**, 278–284.
77. Ito H and Willson C G (1983), 'Chemical amplification in the design of dry developing resist materials', *Polym Eng Sci,* **23**, 1012–1018.
78. Ito H, Ueda M and Schwalm R (1988), 'Highly sensitive thermally developable positive resist systems', *J Vac Sci Technol B,* **6**, 2259–2263.
79. Ito H, Ueda M and Renaldo A F (1989), 'Thermally developable, positive tone oxygen RIE barrier resist for bilayer lithography', *J Electrochem Soc,* **136**, 245–249.
80. Lingnau J, Dammel R and Thesis J (1989), 'Highly sensitive novolak-based x-ray positive resist', *J Polym Eng Sci,* **29**, 874–877.
81. Aoai T, Umehara A, Kamiya A, Matsuda N and Aotani Y (1989), 'Application of silicon polymer as positive photosensitive material', *Polym Eng Sci,* **29**, 887–890.
82. Aoai T, Aotani Y, Umehara A and Kokubo T (1990), 'Application of silylether and silylester polymer for chemical amplification system', *J Photopolym Sci Technol,* **3**, 389–400.
83. Nalamasu O, Tïmko A G, Reichmanis E, Houlihan F M, Novembre A E, Tarascon R, Münzel N and Slater S G (1995), Washington D.C., ACS Symposium Series 614, 4–20, American Chemical Society, Ch.2.
84. MacDonald S *A*, Clecak N J, Wendt H R, Willson C G, Snyder C D, Knors C J, Peyoe N, Maltabes J G, Morrow J, McGuire A E and Holmes S J (1991), 'Airborne chemical contamination of a chemically amplified resist', *Proc SPIE,* **1466**, 2–12.
85. Nalamasu O, Reichmanis E, Hanson J E, Kanga R S, Heimbrook I A, Emerson A B, Baiocchi F A and Vaidya S (1992), 'Effect of post-exposure delay in positive acting chemically amplified resists: an analytical study', *Polym Eng Sci,* **32** 1569–1570.
86. Hashimoto S, Itani T, Yoshino H, Yamana M, Samoto N and Kasama K (1997), 'A study of acid evaporation property in chemically amplified resists', *Proc SPIE,* **3049**, 248–255.
87. Mack C A (1995), *Lithographic Effects of Acid Diffusion in Chemically Amplified Resists*, Washington D.C., ACS Symposium Series 614, 56–68, American Chemical Society.

88. Hinsberg W D, Houle F A, Sanchez M I and Wallraf G M (2001), 'Chemical and physical aspects of the post-exposure baking process used for positive-tone chemically amplified resists', *IBM J Res and Dev,* **45**, 667–682.
89. Wahler, T and Dress P (2008), 'Road to a zero degree total temperature range post exposure bake process', *Proc SPIE,* **7122**, 71220C-1–10.
90. Rothschild M and Ehrlich D J (1988), 'A review of excimer laser projection lithography', *J Vac Sci Technol B,* **6**, 1–17.
91. Shaver D C, Craig D M, Marchi C, Hartney M A and Goodall F (1992), 'Small-filed stepper for 193-nm lithography process development', *Proc SPIE,* **1674**, 766–775.
92. Arnold W A (1999), 'CMOS device fabrication and the evolution of optical lithography exposure tools', *J Microelectronic Eng,* **46**, 7–9.
93. Duffey T, Blumenstock G M, Fleuro V, Pan X J, Newman P, Glatzel H, Watson T, Emmeye J, Kuschnereit R and Weiglb B (2001), 'Next generation 193 nm laser for sub 400 nm Lithography', *Proc SPIE,* **4346**, 1202–1209.
94. Kunz R R, Hartney, M A, Horn M W, Keast C L, Rothschild M and Shaver D C (1993), 'Resist processes for ArF excimer laser lithography', *J Photopolymer Sci Technol,* **4**, 473–490.
95. Allen R D, Conley W E and Kunz R R (1997), Deep-UV resist technology: the evolution of materials and processes for 250-nm and beyond. *Handbook of Microlithography, Micromaching and Microfabrication*, Vol 1. Microlithography, ed. P Rai-Choudhury, Washington D.C. and London, UK SPIE Optical Engineering Press and the Institution of Electrical Engineers, Ch.4.
96. Hinsberg W D and Wallraff G M (2005), *Lithographic Resists Kirk-Othmer Encyclopedia of Chemical Technology,* Hoboken, NJ, USA, John Wiley and Sons, Inc.
97. Schneider J, Greiner A, Lim C- T, Temchenko V, Braun F, Kaiser D, Hauck T, Meusel I, Burmeister D, Loehr S, Volkland S, Bauch A, Kirbach H, Sarlette D and Thiede K (2010), 'Methods and challenges to existing dry 193 nm medium NA lithography beyond 90 nm', *Proc SPIE,* **7640**, 76403G-1–10.
98. Chappell J (2004), 'ASML bridges wet and dry litho, boosts NA', *Electronic News,* **50**(17).
99. Rothschild M, Bloomstein T M, Kunz R R, Liberman V, Switkes M, Palmacci, S T, Sedlacek, J H C, Hardy D and Grenville A (2004), 'Liquid immersion lithography: why, how, and when?', *J Vac Sci Technol, B* **22**, 2877–2881.
100. Gil D, Brunner T A, Fonseca C, Seong N, Streefkerk B, Wagner C and Stavenga M (2004), 'Immersion Lithography: new opportunities for semiconductor manufacturing', *J. Vac Sci Technol B,* **22**, 3431–3438.
101. Bae Y C, Liu Y, Cardolaccia T, Spizuoco K, Bell R, Joesten L, Pikon A, Reilly M, Ablaza S, Trefonas P and Barclay G G (2009), 'Advanced patterning solutions based on double exposure: double patterning and beyond', *Proc SPIE,* **7520**, 75201G-1–9.
102. Wei Y and Brainard R L (2009), *Advanced Processes for 193-nm Immersion Lithography,* Bellingham, WA, SPIE.
103. Kunz R R, Allen R D, Hinsberg W D and Wallraf G M (1993), 'Acid-catalyzed single-layer resists for ArF lithography', *Proc SPIE,* **1925**, 167–175.
104. Allen R D, Wallraff G M, Hinsberg W D, Simpson L L and Kunz R R (1994), *Methacrylate Terpolymer Approach in the Design of a Family of Chemically Amplified Positive Resists,* Washington D.C., ACS Symposium Series 537, 165–177, L.F. Thompson, C.G. Willson eds, American Chemical Society, Ch 11.

105. Allen R D, Hinsberg W D, Wallraf G M and Simpson L L (1991), 'High performance acrylic polymers for chemically amplified photoresist applications', *J Vac Sci Tech B,* **9**, 3357–3361.
106. Allen R D, Wallraff G M, Hinsberg W D, Conley W E and Kunz R R (1993), 'Designing high performance KrF and ArF single layer resists with methacrylate polymers', *J Photopolym Sci Tech,* **6**, 575–592.
107. Kaimoto Y, Nozaki K, Takechi S and Abe N (1992), 'Alicyclic polymer for ArF and KrF excimer resist based on chemical amplification', *Proc SPIE,* **1672**, 66–73.
108. Nozaki K, Kaimoto Y, Takahashi M, Takechi S and Abe N (1994), 'Molecular design and synthesis of 3-oxocyclohexyl methacrylate for ArF and KrF excimer laser resist', *Chem Mater,* **6**, 1492–1498.
109. Ohnishi Y (1981), 'Poly(vinylnaphthalene) and its derivatives as e-beam negative resists', *J Vac Sci Tech,* **19**, 1136–1140.
110. Kunz R R, Palmateer C S, Forte A R, Allen R D, Wallraff G M, DiPietro R A and Hofer D C (1996), 'Limits to etch resistance for 193-nm single-layer-resists', *Proc SPIE,* **2724**, 365–376.
111. Wallow, Brock P, DiPietro R, Allen R, Opitz J, Sooriyakumaran R, Hofer D, Meute J, Byers J, Rich G, McCallum M and Schuetze S (1998), 'Reactive ion etching of 193-nm resist candidates: current platforms and future requirements', *Proc SPIE,* **3333**, 92–101.
112. Shida N, Oshirogouchi T, Asakawa K and Nakasa N (1996), 'Novel ArF excimer laser resists based on menthyl methacrylate terpolymer', *J Photopolym Sci Techn,* **9**, 457–464.
113. Nakano K, Maeda K, Iwasa S, Ohfuji T and Hasegawa E (1995), 'Positive chemically amplified resist for ArF excimer laser lithography composed of a novel transparent photoacid generator and an alicyclic terpolymer', *Proc SPIE,* **2438**, 433–444.
114. Allen R D, Sooriyakumaran R, Opitz J, Wallraff G M, DiPietro R A, Breyta G and Hofer D C (1996), 'Protecting groups for 193-nm photoresists', *Proc SPIE,* **2724**, 334–343.
115. Nozaki K and Yano E (2002), 'High-performance resist materials for ArF excimer laser and electron beam lithography', *Fujitsu Sci Tech J,* **28**, 3–12.
116. Wallow T I, Houlihan F M, Nalamasu O, Chandross E, Neenan T X and Reichmanis E (1996), 'Evaluation of cycloolefin-maleic anhydride alternating copolymers as single-layer photoresists for 193 nm photolithography', *SPIE,* **2724**, 355–364.
117. Reichmanis E, Nalamasu O, Houlihan, F M, Wallow T I, Timko A G, Cirelli R, Dabbagh G, Hutton R S and Novembre A E (1997), 'Resist design concepts for 193 nm lithography: opportunities for innovation and invention', *J Vac Sci Technol B*, **15**, 2528–2533.
118. Houlihan F M, Wallow T, Timko A, Neria E, Hutton R, Cirelli R, Nalamasu O and Reichmanis E (1997), 'Recent advances in 193 nm single-layer photoresists based on alternating copolymers of cyclo-olefins', *Proc SPIE,* **3049**, 84–92.
119. Bloomstein T M, Horn M W, Rothschild M, Kunz R R, Palmacci S T and Goodman R B (1997), 'Lithography with 157 nm lasers', *J Vac Sci Technol B,* **15**, 2112–2117.
120. Kunz R R (1999), 'Outlook for 157 nm resist design', *J Vac Sci Technol B,* **17**, 3267–3273.

121. Ito H, Seehof N, Sato R, Nakayama T and Ueda M (1998), *Synthesis and Evaluation of Alicyclic Backbone Polymers for 193 nm Lithography,* Washington, D.C., ACS Symposium Series 706, Ito, H. Reichmanis, E., Nalamasu, O., Ueda H. eds, 208–223, American Chemical Society.
122. Patel K, Lawson M, Varanasi P, Medeiros D, Wallraff G, Brock P, DiPietro R, Nishimura Y, Chiba T and Selzak M (2004), 'IBM-JSR Negative tone resist: Polymer design, material properties and lithographic performance', *Proc SPIE,* **5376**, 94–113.
123. Ito H Truong H D, Allen R D, Li W, Varanasi P R, Chen, K-J, Khojasteh M, Huang, W S, Burns S D and Pfeiffer D (2006), 'ArF excimer laser resists based on fluoroalcohol', *Polym Adv Technol,* **17**, 104–115.
124. Przybilla K, Roeschert H and Pawloski G (1992), 'Hexafluoroacetone in resist chemistry: a versatile new concept for materials for deep UV lithography', *Proc SPIE,* **1672**, 500–513.
125. Wallraff G M and Hinsberg W D (1999), 'Lithographic imaging techniques for the formation of nanoscopic features', *Chem Rev,* **99**, 1801–1822.
126. Lin B J (2002), 'The k3 coefficient in nonparaxial λ/NA scaling equations for resolution, depth of focus', *J Microlith Microfab Microsyst,* **1**, 7–12.
127. Switkes M and Rothschild M J (2002), 'Resolution enhancement of 157 nm lithography by liquid immersion', *Proc SPIE,* **4691**, 459–465.
128. Hirayama T, Takasu R, Sato M, Wakiya K, Yoshida M and Tamura K (2008), 'Material for forming resist protecting film for use in liquid immersion lithography process, composite film, and method for forming resist pattern', US Patent 7371510.
129. Sundberg L K, Sanders D P, Sooriyakumaran R, Brock P J and Allen R D (2007), 'Contact angles & structure/surface property relationships of immersion materials', *Proc SPIE,* **6519**, 65191Q-1–9.
130. Dammel R R, Houlihan F M, Sakamuri R, Rentkiewitz D and Romano A (2004), '193 nm immersion lithography-taking the plunge', *J Photopolym Sci Tech,* **17**, 587–602.
131. Sanders D P (2010), 'Advances in patterning materials for 193 nm immersion lithography', *Chem Rev,* **110**, 321–360.
132. Ando T, Ohmori K, Maemori S, Takayama T, Ishizuka K, Yoshida M, Hirano T, Yokoya J, Nakano K, Fujiwara T and Owa S (2006), 'Defect studies of resist process for 193 nm immersion lithography', *Proc SPIE,* **6153**, 615309, 1–8.
133. Shirota N, Takebe Y, Sasaki T, Yokokoji O, Toriumi M and Masuhara H (2006), 'Development of fluoropolymer for 193nm immersion lithography', *Proc SPIE,* **6153**, 615324, 1–10.
134. Sanders D P, Sundberg L K, Sooriyakumaran R, Brock P J, DiPietro RA, Truong H D, Miller D C, Lawson M C and Allen R D (2007), 'Fluoro-alcohol materials with tailored interfacial properties for immersion lithography', *Proc SPIE,* **6519**, 651904–1-12.
135. Yamashita T, Ishkawa T, Morita M, Kanemura T and Aoyama H (2008), 'Synthesis of novel α-fluoroacrylates and related polymers for immersion lithography', *Proc SPIE,* **6923**, 69231Z, 1–7.
136. Wei Y, Petrillo K, Brandl S, Goodwin F, Benson P, Housley R and Okoroanyanwu U (2006), 'Selection and evaluation of developer-soluble topcoat for 193 nm immersion lithography', *Proc SPIE,* **6153**, 615306, 1–12.

137. Shedd T A, Schuetter S D, Nellis G F and Van Peski C K (2006), 'Experimental characterization of the receding meniscus under conditions associated with immersion lithography', *Proc SPIE,* **6154**, 61540R, 1–11.
138. Lee S H, Kim J W, Kim J W, Oh S K, Park C S, Lee J Y, Kim S S, Lee J W, Kim D, Kim J, Ban K D, Bok C K and Moon S C (2007), 'Polymer structure modifications for immersion leaching control', *Proc SPIE,* **6519**, 651925, 1–9.
139. Tsuji H, Yoshida M, Ishizuka K, Hirano T, Endo K and Sato M (2005), 'Resist development status for immersion lithography', *Proc SPIE,* **5753**, 102–108.
140. Kanda H and Kanna S (2007), *Positive resist composition, resin used for the positive resist composition, compound used for synthesis of the resin and pattern forming method using the positive resist composition.* US Patent Application 0134588 A1. 2007.
141. Allen R, Brock P, Kusumoto S, Nishimura Y, Sanders D P, Slezak M S, Sooriyakumaran R, Sundberg L K, Trung H and Wallraff G M (2007). *Self-topcoating resist for photolithography.* US Patent Application 0254235 A1. 2007.
142. Wakamatsu G, Anno Y, Hori M, Kakizawa T, Mita M,Hoshiko K, Shioya T, Fujiwara K, Kusumoto S, Yamaguchi Y and Shimokawa T (2009), 'Double patterning process with freezing technique', *Proc SPIE,* **7273**, 72730B-1–8.
143. Chen K J, Huang W S, Li W K and Varanasi P R (2008), 'Resist freezing process for double exposure lithography', *Proc SPIE,* **6923**, 69230G-1–10.
144. Reilly M, Bae Y C and Vohra V (2010), 'Evolution of thermal cure resist for double patterning applications', *Proc SPIE,* **7639**, 76392W-1–7.
145. Maenhoudt M, Versluijs J, Struyf H, Van Olmen J and Van Hove M (2005), 'Double patterning scheme for sub-0.25 k1 single damascene structures at NA = 0.75, λ = 193 nm', *Proc SPIE,* **5754**, 1508–1518.
146. Ooki H, Coon P, Owa S, Sei T and Okamoto K (1997), 'Experimental study on non-linear multiple exposure method', *Proc SPIE,* **3051**, 85–93.
147. Fujisawa T, Anno Y, Hori M, Wakamatsu G, Mita M, Ito K, Tanaka H, Hoshiko K, Shioya T, Goto K, Ogawa Y, Takikawa H, Kozuma Y, Fujiwara K, Sugiura M, Yamaguchi Y and Shimokawa T (2010), 'Simplified double patterning process with non-topcoat self-freezing resist', *Proc SPIE,* **7639**, 76392Y-1–8.
148. Carlson A and Liu T -J K (2008), 'Negative and iterated spacer lithography processes for low variability and ultra-dense integration', *Proc SPIE,* **6924**, 69240B-1–9.
149. Willson C G and Bernard J (2008), 'The future of lithography: SEMATECH litho forum 2008', *ACS Nano,* **2**, 1323–1328.
150. Lee S, Byers J, Jen K, Zimmerman P, Rice B, Turro N J and Willson C G (2008), 'An analysis of double exposure lithography options', *Proc SPIE,* **6924**, 69242A-1–12.
151. Yasin S, Hasko D G and Ahmed H (2002), 'Comparison of MIBK/IPA and water/IPA as PMMA developers for electron beam nanolithography', *Microelectron Eng,* **61–62**, 745–753.
152. Barbee T W, Mrowka S and Hettrick M C (1985), 'Molybdenum-silicon multilayer mirrors for the extreme ultraviolet', *Appl Opt,* **24**, 883–886.
153. Kinoshita H, Kurihara K, Ishii Y and Torii Y (1989), 'Soft x-ray reduction lithography using multilayer mirrors', *J Vac Sci Tech B,* **7**, 1648–1651.
154. Gupta A, Liang R, Tsay F D and Moacanin J (1980), 'Characterization of a dissociative excited state in the solid state: photochemistry of poly(methylmethacrylate). photochemical processes in polymeric systems', *Macromolecules,* **13**, 1696–1700.

155. Gronheid R, Solak H H, Ekinci Y, Jouve A and Van Roey F (2006), 'Characterization of extreme ultraviolet resists with interference lithography', *Microelectron Eng,* **83**, 1103–1106.
156. Rao V, Cobb J L, Henderson C C, Okoroanyanwu U, Bozman D R, Mangat P J S, Brainard R L, Mackevich J F (1999), 'Ultrathin photoresists for EUV lithography', *Proc SPIE,* **3676**, 615–626.
157. Golovkina V N, Nealey P F, Cerrina F, Taylor J W, Solak H H, David C and Gobrecht J (2004), 'Exploring the ultimate resolution of positive-tone chemically amplified resists: 26 nm dense lines using extreme ultraviolet interference lithography', *J Vac Sci Tech B,* **22**, 99–103.
158. Wallraff G M, Medeiros D R, Larson C E, Sanchez M, Petrillo K, Huang W-S, Rettner C, Davis B W, Sundberg L, Hinsberg W D, Houle F A, Hoffnagle J A, Goldfarb D, Temple K, Bucchignano J (2005), 'Studies of acid diffusion in low Ea chemically amplified photoresists', *Proc SPIE,* **5753**, 309–318.
159. Naulleau P, Cain J P, Anderson E, Dean K, Denham P, Goldberg K A, Hoef B and Jackson K (2005), 'Characterization of the synchrotron-based 0.3 numerical aperture extreme ultraviolet microexposure tool at the advanced light source', *J Vac Sci Tech B,* **23**, 2840–2843.
160. Dai J and Ober C K (2003), 'Synthesis and evaluation of novel organoelement resists for EUV lithography', *Proc SPIE,* **5039**, 1164–1172.
161. Junarsa I, Stoykovich M P, Nealey P F, Ma Y, Cerrina F and Solak H H (2005), 'Hydrogen silsesquioxane as a high resolution negative-tone resist for extreme ultraviolet lithography', *J Vac Sci Tech B,* **23**, 138–143.
162. Watanabe T, Hada H, Lee S Y, Kinoshita H, Hamamoto K and Komano H (2005), 'Development of fast-photospeed chemically amplified resist in extreme ultraviolet lithography', *Jpn J Appl Phys, Part 1*, **44**, 5866–5870.
163. Hirose R, Kozawa T, Tagawa S, Kai T and Shimokawa T (2008), 'Dependence of acid generation efficiency on acid molecular structure and concentration of acid generator in chemically amplified EUV resist', *Proc SPIE,* **6923**, 69232A-1–9.
164. Gonsalves K E, Thiyagarajan M and Dean K (2005), 'Newly developed polymer bound photoacid generator resist for sub-100 nm pattern by EUV lithography', *Proc SPIE,* **5753**, 771–777.
165. McCord M A (1997), 'Electron beam lithography for 0.13 um manufacturing', *J Vac Sci Technol B*, **15**, 2125–2129.
166. Skinner J G, Groves T G, Novembre A E, Pfeiffer H and Sing R (1997), Photomask fabrication procedures and limitations. In: *Handbook of Microlithography, Micromaching and Microfabrication,* Vol 1. Microlithography, P. Rai-Choudhury, ed., UK SPIE Optical Engineering Press and the Institution of Electrical Engineers, Ch.5.
167. Rizvi S A (2005), Mask processing. In: *Handbook of Photomask Manufacturing Technology,* S Rizvi ed., Boca Raton, CRC Press, Ch.17.
168. Babin S (1997), 'Measurement of resist heating in photomask fabrication', *J Vac Sci Technol B,* **15**, 2209–2213.
169. Shamoun B (2001), 'Photomask patterning: the influence of substrate bulk heating on placement accuracy', *Microelectron Eng,* **57–58**, 447–452.
170. Harroitt L R (1997), 'Scattering with angular limitation projection electron beam lithography for suboptical lithography', *J Vac Sci Technol B*, **15**, 2130–2135.
171. Steenbrink S W H K, Kampherbeek B J, Wieland M J, Chen J H, Chang S M, Pas M, Kretz J, Hohle C, van Steenwinckel D, Manakli S, Le-Denmat J and Pain

L (2008), 'High throughput maskless lithography: low voltage versus high voltage', *Proc SPIE,* **6921**, 6921LT-1–10.

172. Icard B, Rio D, Veltman P, Kampherbeek B, Constancias C and Pain L (2009), 'Development of resist process for 5 kV multi-beam technology', *Proc SPIE,* **7271**, 72710R-1–12.
173. Novembre A and Bowmer T N (1984), A novel technique for determining radiation chemical yields of negative electron-beam resists. In: *Materials for Microlithography* L.F. Thompson, C.G. Willson, J.M.J. Fréchet, eds,. Washington D.C., ACS Symposium Series 266, American Chemical Society, pp. 241–254.
174. Charlesby A (1960), *Atomic Radiation and Polymers.* In: N.G.Gaylord and D S. Ballantine, eds, New York, Pergamon Press.
175. Heidenreich R D, Thompson L F, Feit E D and Melliar C M (1973), 'Fundamental aspects of electron beam lithography.I. Depth-dose response of polymeric electron beam resists', *J Appl Phys,* **44**, 403.
176. Willson C G (1983), Introduction to Microlithography. In: L.F. Thompson, C.G. Willson and M.J.Bowden, eds, Washington, D.C., ACS Symposium Series 219, American Chemical Society, Ch.3.
177. Pittman C U Jr, Iqbal M, Chen C Y and Helbert N (1978),'Radiation degradation of poly(α-hyrdroxyisobutyric acid) and poly(glycolic ester)', *J Polymer Sci, Polymer Chem,* **16**, 2721–2724.
178. Hiraoka H (1977), 'Radiation chemistry of poly(methacrylates)', *IBM J Res Dev,* **21**, 121–130.
179. Greeneich J S (1975), 'Developer characteristics of poly (methylmethacrylate) electron resist', *J Electrochem Soc,* **122**, 970–976.
180. Delaire J A, Lagarde M, Broussoux D and Dubois J C (1990), 'Effects of molecular weights and polydispersity on the properties of poly(trifluoroethyl methacrylate) as a positive x-ray and electron resist', *J Vac Sci Technol B*, **8**, 33–38.
181. Zhou J L and Cartwright A N (2006), 'Reduction of metal linewidth through a combination of low temperature and ultrasonic development of poly(methylmethacrylate) using electron beam lithography', *Proc SPIE,* **6327**, 6370O-1–11.
182. Küpper Dav, Küpper Dan, Wahlbrink T, Bolten J, Lemme M C, Georgiev M and Kurz, H (2006), 'Megasonic-assisted development of nanostructures', *J Vac Sci Technol B*, **24**, 1827–1832.
183. Rathsack B, Medeiros D and Willson C G (2005), Resists for Mask Making. In: *Handbook of Photomask Manufacturing Technology,* S. Rizvi, ed, Boca Raton, FL., CRC Press, Ch.15.
184. Bowden M J (1985), A perspective on resist materials for fine-line lithography. In: *Materials for Microlithography,* L.F. Thompson, C.G. Willson and J.M.J. Fréchet, eds, Washington D C, ACS Symposium Series 266, American Chemical Society, pp 39–124.
185. Bowden M J and Thompson L F (1973), 'Electron irradiation of poly(olefin sulfones); application to electron beam resists', *J Appl Polymer Sci,* **17**, 3211–3221.
186. Bowden M J, Thompson L F, Farenholtz S R and Dorries E M (1981), 'A sensitive novolac-based positive electron resist', *J Electrochem Soc*, 128, 1304–1312.
187. Mueller M, Komarov S and Baik K H (2002), 'Dry etching of chrome for photomasks for the 100-nm technology using chemically amplified resist', *Proc SPIE,* **4754**, 350–360.

188. Pease R F and Chou S Y (2008), 'Lithography and other patterning techniques for future electronics', *Proc IEEE,* **96**, 248–270.
189. Sauer C A, Dean R L, Etsuya M, Tan Z C H, Smith B W, Ewbank D E, Duttagupta S P and Rudack A (1995), 'Evaluation of commercial and experimental resist materials for use in MEBES mask making', *Proc SPIE,* **2621**, 52–61.
190. Brown J R and O'Donnell J (1970), 'The degradation of poly(butene-1-sulfone) during *Υ* irradiation', *Macromolecules,* **3**, 265–267.
191. Reichmanis E and Thompson L F (1987), 'Polymer materials for microlithography', *Ann Rev Mater Sci,* **17**, 235–271.
192. Dean R, Lem H, Sauer C and Chang H (1994), 'PBS resist profile studies for submicron mask lithography', *Proc SPIE*, **2322**, 102–113.
193. Kim M Y, Lee J H, Yoon Y J and Choi B Y (2000), '150-nm mask fabrication using thin ZEP 7000 resist, Ghost, and dry etch for MEBES 5000 pattern generator', *Proc SPIE,* **4066**, 243–251.
194. Medeiros D R, Aviram A, Guarnieri C R, Huang W-S, Kwong, Magg C K, Mahorowala A P, Moreau W M, Petrillo K E and Angelopoulos M (2000), 'Recent progress in electron-beam resists for advanced mask-making,' *IBM J Res Dev,* **45**, 639–650.
195. Coleman T and Lu A M (1997), 'Comparison of EBR-900 MI and ZEP 7000 with plasma-etch processing for MEBES 4500S,' *Proc SPIE*, **3236** 397–404.
196. Tan Z C H, Le P and Coleman T (2001), 'Potential of DNQ/novolac and chemically amplified resists for 100 nm device generation maskmaking', *Microelectron Eng,* **57–58**, 531–538.
197. Ito H, Pederson L A, MacDonald S A, Cheng Y Y, Lyerla J R and Willson C G (1988), 'A sensitive, etch resistant, positive tone e-beam resist system,' *J Electrochem Soc,* **135**, 1504–1508.
198. Irmscher M, Beyer D, Butschke J, Constantine C, Hoffmann T, Koepernik C, Krauss C, Leibold B, Florian Letzkus, Mueller D, Springer R and Voehringer P (2002), 'Comparative evaluation of e-beam sensitive chemically amplified resists for mask making', *Proc SPIE,* **4754**, 176–187.
199. Tokunaga K, Koba F, Miyasaka M, Takaishi Y, Noda K, Yamashita H, Nakajima K and Nozue H (2000), '*EB Projection Lithography for 60–80 nm ULSl Fabrication*', Symposium on VLSI Technol. Digest of Tech. papers. 54–55.
200. Harriott L R, Berger S D, Liddle J A, Watson G P and Mkrtchyan M M (1995), 'Space charge effects in projection charged particle lithography systems', *J Vac Sci Technol B*, **13**, 2404–2408.
201. Sato M, Ocola L E, Novembre A E, Ohmori K, Ishikawa K, Katsumata K and Nakayama T (1999), 'Characteristics for negative and positive tone resists with direct write electron beam and SCALPEL exposure systems', *J Vac Sci Technol B*, **17**, 2873–2877.
202. Stanton S T, Liddle J A, Gallatin G M, Kim B and Engelstad R (1999), 'Initial wafer heating analysis for a SCALPEL lithography system', *Microlectron Eng,* **46**, 235–238.
203. Liddle J A, Gallatin G M and Ocola L E (2003), 'Resist requirements and limitations for nanoscale electron-beam patterning', *Mat Res Soc Symp Proc,* **739**, H1.5.1-H1.5.11.
204. Ito H, Breyta G, Conley W, Hagerty P, Thackeray J, Holmes S, Nunes R, Fenzel-Alexander D, Dipietro R and Hofer D (1996), 'Lithographic feasibility of ESCAP beyond quarter micron', *J Photopolym Sci Technol,* **9**, 557–572.

205. Harriott L R, Berger S D, Biddick C, Blakey M I, Bowler S W, Brady K, Camarda R M, Connelly W F, Crorken A, Custy J, DeMarco R, Farrow R C, Felker J A, Fetter L, Freeman R, Hopkins L, Huggins H A, Knurek C S, Kraus J S, Liddle J A, Mkrtychan M, Novembre A E, Peabody M L, Tarascon R G, Wade H H, Waskiewicz W K, Watson G P, Werder K S and Windt D (1997), 'The SCALPEL proof of concept system', *Microlectron Eng,* **35**, 477–480.
206. Thompson L F, Ballantyne J P and Feit E D (1975), 'Molecular parameters and lithographic performance of poly(glycidyl methacrylate-co-ethyl acrylate): a negative electron resist', *J Vac Sci Technol,* **12**, 128–1283.
207. Tan C H Z and Georgia S S (1983), 'Evaluation of poly(allyl methacrylate)-co-(hydroxyethyl methacrylate) as negative electron-beam resist', *Polym Eng Sci*, **23**, 963–967.
208. Taniguchi Y, Hatano Y, Shiraishi H, Horigome S, Nonogaki S and Naraoka K (1979), 'PGMA as a high resolution high sensitivity negative electron beam resist', *Jpn J Appl Phys,* **28**, 1143–1148.
209. Hoyle C E and Kinstle J F (1989), *Radiation Curing of Polymeric Materials*, Washington D C, ACS Symposium Series 417, American Chemical Society.
210. Novembre A E and Bowden M J (1983), 'Effect of varying the composition of copolymers of glycidyl methacrylate and 3-chlorostyrene (GMC) on electron lithographic performance', *Polymer Eng Sci,* **23**, 975–979.
211. Itaya K and Shibayama K (1982), 'High resolution electron beam negative resist with very narrow molecular weight distribution', *J Electrochem Soc,* **129**, 663–665.
212. Helbert J N, Iafrate G J, Pittman C U Jr and Lai J H (1980), 'Effect of chemical composition upon the radiation and electron beam resist behaviors of vinyl polymers', *Polymer Eng Sci,* **20**, 1077–1081.
213. Jones R G, Miller Tate R G and Brambley D R (1991), 'Radiation chemical yields and lithographic performance of electron-beam resists based on poly(methylstyrene-*co*-chlorostyrene)', *J Mater Chem,* **1**, 401–407.
214. O'Donnell J H (1991), Chemistry of radiation degradation of polymer. In: *Radiation Effects on Polymers*, R.L. Clough and S. W. Shalaby, eds, Washington D.C., ACS Symposium Series 475, American Chemical Society. Ch.24.
215. Charlesby A and Pinner S H (1959), 'Analysis of the solubility behaviour of irradiated polyethylene and other polymers', *Proc R Soc London,* Ser. *A*., **249**, 367–386.
216. Güven O (1989)*, Crosslinking and Scission in polymers*, *Proc. NATO ASI*, 202, Dordrecht, Netherlands, Kluwer Academic Publishers.
217. Feit E D and Stillwagon L E (1980), 'Electron-Beam lithography of chlorinated polystyrenes with narrow molecular weight distributions', *Polymer Eng Sci,* **20**, 1058–1063.
218. Hartney M A (1989), 'Analysis of radiation chemical yields of chlorinated polystyrene derivatives', *J Appl Polymer Sci,* **37**, 695–705.
219. Brambley D R, Jones R G, Matsubayashi Y and Miller Tate P (1990),'Structural control and optimization of chlorinated styrene based electron resists', *J Vac Sci Technol B,* **8**, 1412–1417.
220. Choong H S and Kahn F J (1981), 'Molecular parameters and lithographic performance of poly(chloromethlystyrene)- a high performance negative electron resist', *J Vac Sci Technol B*, **19**, 1121–1126.

221. Hartney M A, Tarascon R G and Novembre A E (1985), 'Lithographic evaluation and processing of chlorinated polymethylstyrene', *J Vac Sci Technol B*, **3**, 360–366.
222. Taylor G N and Wolf T M (1980), 'Oxygen plasma removal of thin polymer films', *Polymer Eng Sci,* **20**, 1087–1092.
223. Tedesco S, Pierrat C, Lamure J M, Sourd C, Martin J and Guibert J C (1990), 'Dry etching for high resolution maskmaking', *Proc SPIE,* **1264**, 1.
224. Kataoka M, Kanetsuki S, Tamura K, Yamamoto K and Asano M (1994), 'A new novolac-based positive EB resist, EBR-900 M-1', *Proc SPIE,* **2254**, 47.
225. Novembre A E, Mixon D A, Pierrat C, Knurek C and Stohl M (1993), 'Dry etch patterning of chrome on glass optical masks using P(SI-CMS)', *Proc SPIE,* **2087**, 50–56.
226. Novembre A E, Reichmanis E and Davis M (1986), 'Preparation and lithographic properties of poly(trimethylsilylmethyl methacrylate-co-chloromethyl styrene)', *Proc SPIE,* **631**, 14–21.
227. Lee K Y, LaBianca N, Rishton S A, Zolgharnain S, Gelorme J D, Shaw J and Chang T H P (1995), 'Micromaching applications of a high resolution ultrathick photoresist', *J Vac Sci Technol B*, **13**, 3012–3016.
228. Kouba J, Engelke R, Bednarzik M, Ahrens G, Scheunemann H-U, Gruetzner G, Loechel B, Miller H and Haase D (2007), 'SU-8 promising resist for advanced direct LIGA applications for high aspect ratio mechanical microparts', *Microsyst Technol,* **13**, 311–317.
229. Shaw J M, Gelorme J D, LaBianca N C, Conley W E and Holmes S J (1997), 'Negative photoresists for optical lithography', *IBM J Res Develop,* **41**, 81–94.
230. Boganov A L (2000), 'Use of SU-8 negative photoresist for optical mask manufacturing', *Proc SPIE,* **3999**, 1215–1225.
231. Aktary M, Jensen M O, Westra K L, Brett M J and Freeman M R (2003), 'High-resolution pattern generation using the epoxy novolack SU-8 2000 resist by electron beam lithography', *J Vac Sci Technol B,* **21**, L5–L7.
232. Bilenberg B, Jacobsen S, Schmidt M S, Skjolding L H D, Shi P, Bøggild P, Tegenfeldt J O and Kristensen A (2006), 'High resolution 100 kV electron beam lithography in SU-8', *Microelectron Eng,* **83**, 1609–1612.
233. Saint-Pol J, Landis S, Gourgon C, Tedesco S, Hanawa R, Suetsugu M, Akita M and Yamamoto S (2003), 'Negative tone chemically amplified resist formulation optimizations for ultra high-resolution lithography', *Microelectron Eng,* **67–68**, 274–282.
234. Claßen A, Kuhn S, Straka J and Forchel A (1992), 'High voltage electron beam lithography of the resolution limits of SAL 601 negative resist', *Microelectron Eng*, **17**, 21–24.
235. Ocola L E, Tennant D, Timp G and Novembre A (1999), 'Lithography for sub-60 nm resist nanostructures', *J Vac Sci Technol B*, **17**, 3164–3167.
236. Novembre A E (2000), Presentation materials at 4th NGL Sematech workshop, September 25–26, Reston, Virginia.
237. Studer V, Pépin A and Chen Y (2002), 'Nanoembossing of thermoplastic polymers for microfluidic applications', *Appl Phys Lett,* **80**, 3614–3616.
238. Van Dodewaard A J, Ketelaars W S M M, Roes R F M, Kwinten J A J, Van Delft F CM J M, Van Run A J, Van Langen-Suurling A K and Romijn J (2000), 'Comparison of negative tone resists NEB22 and UVN30 in e-beam lithography', *Microlectron Eng,* **53**, 461–464.

239. Gutsche D C (2008), *Calixarenes An Introduction,* 2nd ed.' Stoddard J.F. ed., Cambridge UK, Royal Society of Chemistry.
240. Fujita J, Ohnishi Y, Ochiai Y and Matsui S (1996), 'Ultrahigh resolution of calixarene negative resist in electron lithography', *Appl Phys Lett,* **68** 1297–1299.
241. Fujita J, Ohnishi Y, Ochiai Y, Nomura E and Matsui S (1996), 'Nanometer-scale resolution of calixarene negative resist in electron beam lithography', *J Vac Sci Technol B*, **14**, 4272–4276.
242. Sailer H, Ruderisch A, Kern D P and Schurig V (2002), 'Evaluation of calixarene derivatives as high-resolution negative tone electron-beam resists', *J Vac Sci Technol B*, **20**, 2958–2961.
243. Auzelyte V, Langner A and Solak H H (2009), 'Thermal development of a calixarene resist', *J Vac Sci Technol B*, **27**, 2990–2992.
244. Yang C C and Chen W C (2002), 'The structures and properties of hydrogen silsesquioxane (HSQ) films produced by thermal curing', *J Mater Chem,* **12**, 1138–1141.
245. Ro H W and Soles C L (2011), 'Silsesquioxanes in nanoscale patterning applications', *Mater Today,* **14**, 20–33.
246. Namatsu H, Yamaguchi T, Nagase M, Yamazaki K and Kurihara K (1998), 'Nano-patterning of a hydrogen silsequioxane resist with reduced linewidth fluctuations', *Microelectron Eng,* **41/42**, 331–334.
247. Olynick D L, Cord B, Schipotinin A, Ogletree D F and Schuck P J (2010), 'Electron-beam exposure mechanisms in hydrogen silsesquioxane investigated by vibrational spectroscopy and in situ electron-beam-induced desorption', *J Vac Sci Technol B,* **28**, 581–587.
248. Word M J, Adesida I and Berger P R (2003), 'Nanometer-period gratings in hydrogen silsesquioxane fabricated by electron beam lithography', *J Vac Sci Technol B,* **21**, L12–L15.
249. Falco C M, van Delft J M and Weterings J P (2000), 'Hydrogen silsesquioxane/novolak bilayer resist for high aspect ratio nanoscale electron-beam lithography', *J Vac Sci Technol B*, **18**, 3419–3423.
250. Telecky A, Xie P, Stowers J, Grenville A, Smith B and Keszler D A (2010), 'Photopatternable inorganic hardmask', *J Vac Sci Technol B*, **28** C6S19–C6S22.
251. Saifullah M S M, Khan M Z R, Hasko D G, Leong E S P, Neo X L, Goh E T L, Anderson D, Jones G A C and Welland M E (2010) 'Spin-coatable HfO_2 resist for optical and electron beam lithographies', *J Vac Sci Technol B,* **28**, 90–95.
252. Stowers J and Keszler D (2009), 'High resolution, high sensitivity inorganic resists', *Microelectron Eng,* **86**, 730–733.
253. Bhushan B (2007), *Springer Handbook of Nanotechnology*, 2nd ed., New York, Springer.
254. Wang J (2005), 'Tutorial: status of nanoimprint lithography and device applications', *Proc SPIE,* **6013**, 6013021–15.
255. Palmieri F, Adams J, Long B, Heath W, Tsiartas P, Willson C G (2007), 'Design of reversible cross-linkers for step and flash imprint lithography imprint resists', *ACS Nano*, **1**, 307–312.
256. Costner E, Lin M, Jen W and Willson C G (2009), 'Nanoimprint lithography materials development for semiconductor device fabrication', *Ann Rev Mater Res,* **39**, 155–180.

257. Piner R D, Zhu J, Xu F, Hong S and Mirkin C A (1999), 'Dip-pen nanolithography', *Science*, **283**, 661–663.
258. Basnar B and Willner I (2009), 'Dip-pen-nanolithographic patterning of metallic, semiconductor, and metal oxide nanostructures on surface', *Small*, **5**, 28–44.
259. Ginger D S, Zhang H and Mirkin C A (2004), 'The evolution of dip-pen nanolithography', *Angew Chem Int Ed,* **43**, 30–45.
260. Salaita K, Wang Y H, and Fragala J, Vega R A, Liu C and Mirkin C A (2006), 'Massively parallel dip-pen nanolithography with 55000-pen two-dimensional array', *Angew Chem Int Ed,* **45**, 7220–7223.
261. Giam L R and Mirkin C A (2011), 'Cantilever-free scanning probe molecular printing', *Angew Chem Int Ed,* **50**, 7482–7485.
262. Huo F, Zheng Z and Zheng G, Giam L R, Zhung H and Mirkin C A (2008), 'Polymer pen lithography', *Science*, **321**, 1658–1660.
263. Zheng Z, Daniel W L, Giam L R, Huo F, Senesi A J, Zheng G and Mirkin C A (2009), 'Multiplexed protein array enabled by polymer pen lithography: addressing the inking challenge', *Angew Chem Int Ed,* **48**, 7626–7629.
264. Tanaka K, Naito R, Kitada T, Kiba Y, Yamada Y, Kobayashi M and Ichikawa H (2003), 'Improvement of pattern collapse issue by additive added D.I. water rinse process', *Proc SPIE,* **5039**, 1366–1381
265. Jouve A, Simon J and Foucher J, David T, Tortai J-H and Solak H (2005), 'Overcoming pattern collapse of ultra high resolution dense lines obtained with EUV resists', *Proc SPIE,* **5753**, 720–731.
266. Miyahara O, Tanaka K and Wakamizu S, Kitano J and Yamada Y (2004), 'Improvement of pattern collapse issue by additive added D.I. water rinse process 2', *Proc SPIE,* **5376**, 830–841.
267. Hinsberg W, Houle F, Sanchez M, Hoffnagle J, Walraff G, Medeiros D, Gallatin G and Cobb J (2003), 'Extendibility of chemically amplified resist: another brick wall?', *Proc SPIE,* **5039**, 1–14.

8

Directed assembly nanolithography

S. MATSUI, University of Hyogo, Japan, M. TAKENAKA,
Kyoto University, Japan and H. YOSHIDA, Hitachi
Research Laboratory, Japan

DOI: 10.1533/9780857098757.287

Abstract: Directed assembly is a promising nanolithography to produce a complex structure with small feature size. The chapter first discusses block copolymer lithography to realize the full potential of self-assembled structure. Secondly, programmable three-dimensional lithography using focused-ion-beam and electron-beam chemical-vapor-deposition is described.

Key words: directed assembly, directed self-assembly, block copolymer, three-dimensional lithography.

8.1 Introduction

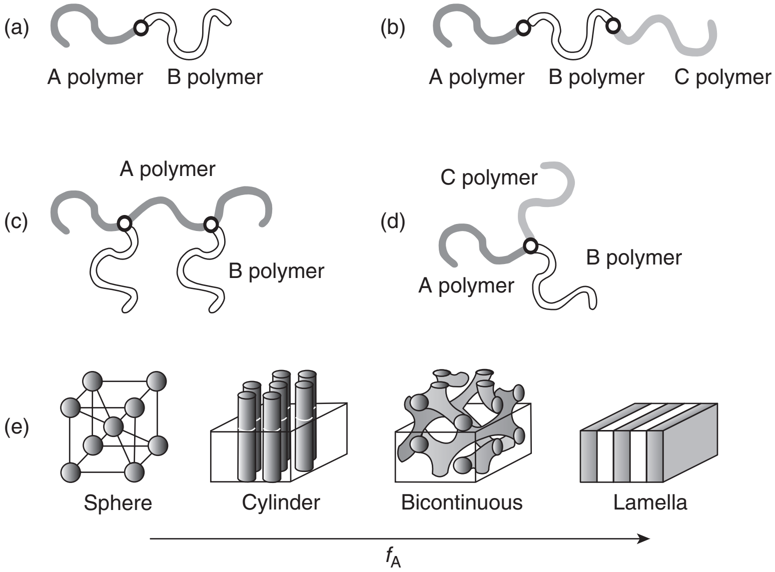
The use of directed assembly for nanolithography has received much attention for organizing materials on the nanoscale into ordered structures and producing complex structures with small feature sizes. These directed assembly structures promise new opportunities for developing miniaturized optical, electronic devices and magnetic devices. So far, self-assembly using block copolymer and nanoparticles, sphere lithography, edge lithography, programmable three-dimensional lithography, and so on have been described as directed assembly lithography. In this chapter, we will first describe block copolymer lithography to realize the full potential of self-assembled structures and to utilize them in more demanding applications such as magnetic storage or semiconductor devices. Next, we will describe programmable three-dimensional lithography using focused-ion-beam (FIB) and electron-beam (EB) chemical-vapor-deposition (CVD).

8.2 Block copolymers in lithography

Block copolymers are promising materials for lithography to attain smaller dimensions, overcoming some of the limitations of current top-down fabrication processes. In this section we will first describe the physics of block copolymer self-assembly in bulk and in thin films, followed by a discussion of the methods to control self-assembly of block copolymer for nanoscale lithography.

8.2.1 Microphase separation of block copolymers in bulk and in thin films

Block copolymers are composed of two or more kinds of chemically distinct polymers linked by a covalent bond. A variety of architectures of block copolymers with significantly small molecular weight distribution, such as diblock, triblock, multi-block, star-block, and graft copolymer have been synthesized by living polymerization techniques. The simplest block copolymer is the linear type AB diblock copolymer, which still can self-assemble into various periodic microdomain structures via microphase separation, as shown in Fig. 8.1 (Hamley1998). The morphologies of diblock copolymers depend on several parameters, such as the volume fraction of one block in diblock copolymer (*f*), the Flory-Huggins interaction parameter between two block chains (χ), which represents the energy cost of two spices mixed homogeneously and is inversely proportional to temperature, and the polymerization index or the total number of monomers per polymer chain of

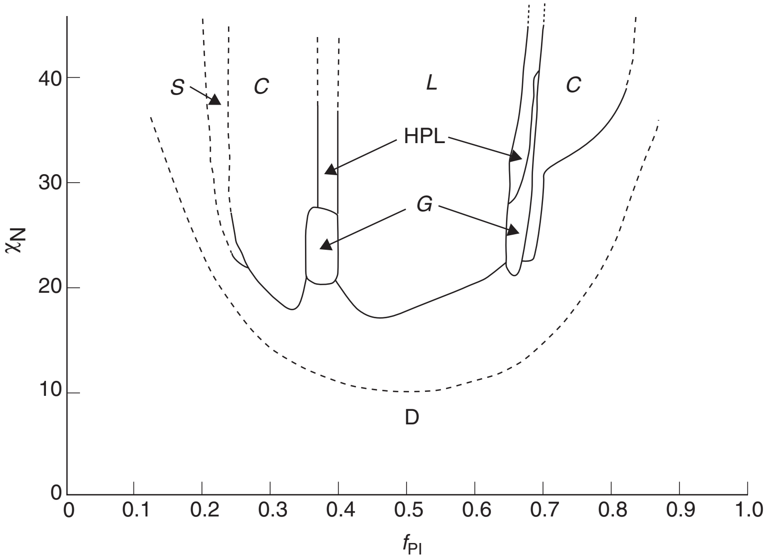


8.1 Depictions of various architectures of block copolymers: (a) Diblock copolymer, (b) ABC linear triblock terpolymer, (c) (AB)A(AB) π-shaped graft copolymer, (d) ABC star triblock terpolymer, and (e) four ordered morphologies in AB diblock copolymer.

the diblock copolymer ($N = N_A + N_B$), where subscripts A and B represent the two distinct block chains and N_i is total number of monomers in the ith chain. The strength of the segregation of diblock copolymer is expressed by the reduced parameter, χN. When χN does not exceed a critical value, $(\chi N)_{\mathrm{ODT}}$ (ODT: order–disorder transition), A and B block chains are mixed in the molecular level (disordered state), which is similar to the case of a blend of A and B homopolymer chains. In contrast, when χN exceeds a critical value, $(\chi N)_{\mathrm{ODT}}$, A and B block chains segregate into A and B rich phases, respectively, (ordered state). Because the chemical junctions of A and B block chains are located at the interface between two domains, the phase separation occurs in microscopic scale, not in macroscopic scale (larger than 1 μm) as is commonly the case for the blend of A and B polymer chains. The dimensions of domains formed by microphase separation are therefore on the order of the radius of the gyration, R_g, of each block chain: ~5–100 nm. For this reason, diblock copolymers form 'microdomain structures' via 'microphase separation,' such as lamellae (L), gyroid (G), hexagonally-packed cylinders (C), and spheres in a body-centered lattice (S) as equilibrium structures depicted in Fig. 8.1e.

The phase diagram of diblock copolymer melts in bulk state has been investigated over several decades, both theoretically and experimentally, and can be controlled by f and χN (Hamley, 1998). Matsen and Schick examined the phase diagram of diblock copolymer by using self-consistent field theory (SCFT) and found four morphologies mentioned above: L, G, C, and S. Khandpur *et al.* (1995) studied the phase diagram of polystyrene-*block*-polyisoprene (SI) diblock copolymer near the ODT experimentally and observed the S, C, L, and G morphologies, as shown in Fig. 8.2 (Hamley, 1998). Additionally, a fifth morphology, F*ddd*, which is an interconnected but uniaxial structure, has been found as an equilibrium phase between L and G phases (Takenaka *et al.,* 2007).

In thin films, where the block copolymers are confined by a film thickness comparable to the dimension of the individual polymer domains, the surface area per volume in thin film is much larger than that in the bulk state (Fasolka and Mayes, 2001; Hamley, 2009). Therefore, the morphologies and orientation of the microdomains strongly depend on the chain conformations of diblock copolymers in confined space and on the interactions at the free (air–film) and substrate (substrate–film) interfaces. If the free and substrate surfaces have a preferential affinity for one component in the block copolymer, the component wets the preferential surface. The existence of preferential affinity by the surfaces induces the microdomains to orient parallel (in-plane) to substrate. On the other hand, when the constituent polymers in diblock copolymers do not show any preferential affinity to the surfaces, i.e. when the surface is 'neutral' to both components of block copolymer, the microdomains prefer to orient perpendicular (out-plane) to

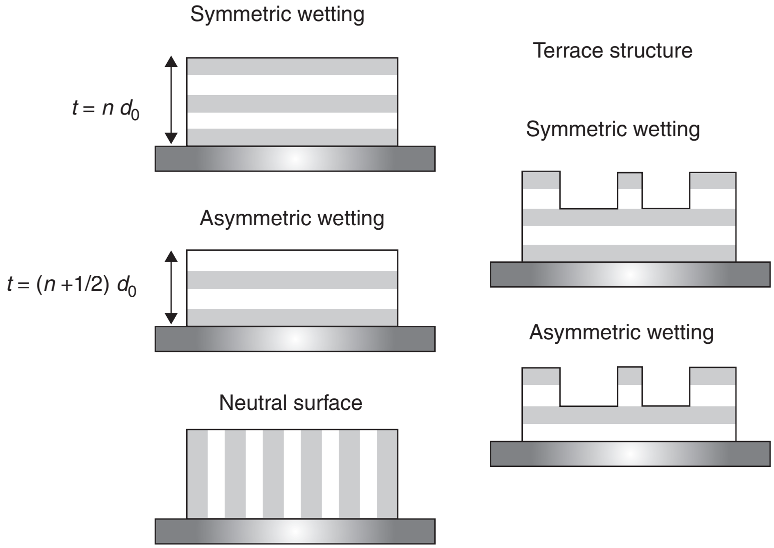


8.2 Experimental phase diagram of polystyrene-*block*-polyisoprene diblock copolymer melts by Khandpur *et al.* (1995), where f_{PI} represents the volume fraction of polyisoprene. Note that *L* is lamellar, *G* is gyroid, H is hexagonally-packed cylinders, *C* is spheres on the body-centered lattice, D is disordered state, and HPL is hexagonally perforated layer.

the substrate to attain an energetically favorable state of the chain conformations (Fasolka and Mayes, 2001).

For in-plane orientation, morphology as well as macroscopic film structure is controlled by commensurability between domain spacing and film thickness, where domain spacing is the distance between nearest neighbor domains. In the case of the thin film of diblock copolymers forming lamellar structures, the preferential affinity quantizes the film thickness, as shown in Fig. 8.3. For symmetric wetting condition, where the same component wets both surfaces, the thickness becomes integer multiples of the domain spacing d_0 or $t = n \cdot d_0$, where t is the film thickness and n is an integer. For asymmetric wetting condition, half-integer thicknesses or $t = (n + 1/2)\ d_0$, is stable. If the thickness of uniform films of the block copolymer does not conform to the discretized values, the films form terrace structures (islands or holes) in the surface of the films. The formation of the terrace is due to the fact that the energetic penalty caused by the terrace structures is less than that caused by the chain stretching or compression in microdomains with unfavored domain spacing.

Similar to the case of the lamellar structure, for the diblock copolymers forming cylindrical structures, the orientation parallel to the surface is induced when one component in diblock copolymers has a preferential



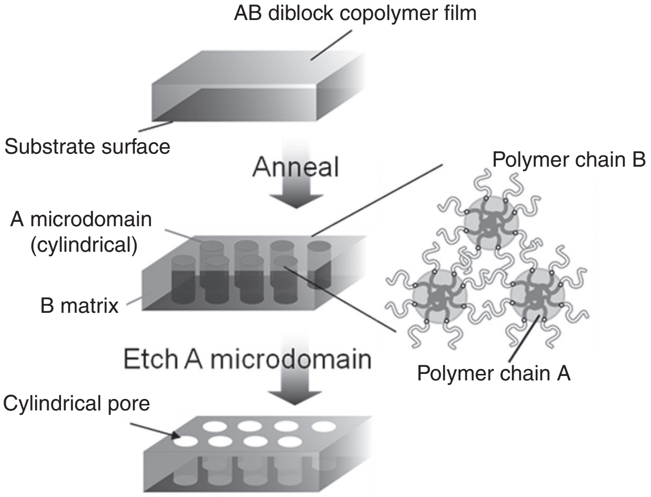
8.3 Orientation of lamellar structures under various interfacial affinities in thin films.

affinity to the substrate or free interfaces. However, since the composition (f) of diblock copolymer forming cylinder structures is asymmetric, the preferential affinity for a given surface induces interesting morphologies. If a surface has a preferential affinity to the minor component with volume fraction being less than in the block copolymers, the minor component covers the surface entirely and a brush layer along the surface is formed (Park *et al.,* 2008a). If there is no preferential affinity by any block at the free surface, a layer of half cylinders is formed at $t = (n + 1/2)\ d_0$ (Harrison *et al.,* 2000). When the terrace structures are formed, perpendicularly oriented cylinders and perforated lamellar structures (Harrison *et al.,* 1998; Zhang *et al.*, 2008) are observed at the edge of the terrace steps.

8.2.2 Block copolymer lithography

Periodic microdomain structures of block copolymers assembled in thin films can be applied as masks for nanoscale lithography. Mansky *et al.* (1995, 1996) first demonstrated that monolayer films of diblock copolymer microdomains could potentially be used as masks on the scale of a few tens of nanometers. Park *et al.* (1997) demonstrated pattern transfer of spherical or cylindrical microdomain structures to substrates by using self-assembled block copolymer thin films as etching masks.

Figure 8.4 illustrates typical process of block copolymer lithography. First, block copolymer thin film is deposited on the substrate by spin coating from

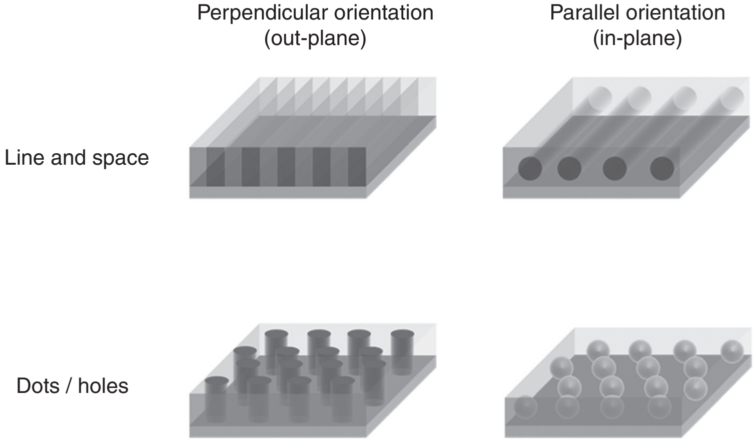


8.4 Schematic illustration demonstrating block copolymer lithography process taking vertically oriented cylindrical microdomain as an example of block copolymer microphase separated structure.

the solution. The structures as spun block copolymers are frozen in the very early stage of microphase separation due to rapid solvent evaporation in the spin-coating process. Therefore, to attain equilibrium microdomain structures, the samples are annealed by thermal treatment or exposure to solvent vapor to apply mobility to the block copolymer chains to promote microphase separation. After forming well-defined microdomain structures, one of the domains in the microphase separated block copolymer is selectively removed by etching. The porous film thus obtained can be used as a mask to etch the substrate or a scaffold to transfer the pattern to various materials.

For a block copolymer film to be able to form a lithographic mask, the following requirements have to be met. First, the constituent blocks of the block copolymer need to display enough etching contrast to selectively remove one of the domains. For example, polystyrene (PS) has higher resistance against O_2 RIE and UV radiation compared to polymethylmethacrylate (PMMA). Therefore, polystyrene-*block*-polymethylmethacrylate (PS-*b*-PMMA) is one of the most commonly applied materials for block copolymer lithography. After forming microdomain structures in thin films on the substrate, PMMA domains can be selectively removed by applying a dry process using O_2 RIE or by a wet process with UV treatment followed by immersion in acetic acid.

Another important point to be considered is the microstructure of the domains to be applied. As illustrated in Fig. 8.5, line and stripe (L&S) patterns can be transferred to the substrate applying either perpendicularly



8.5 Domain structures and their orientation for forming line-and-space and dots/holes patterns.

oriented lamellae or a single layer of parallel-oriented cylinders. Similarly, dot or hole patterns can be obtained by applying perpendicularly oriented cylinders or a single layer of spheres. As can be easily imagined, the perpendicularly oriented domains are preferable as masks for pattern transfer, because the cross-section of the domains is straight and oriented perpendicular to the substrate. Moreover, as discussed in Section 8.2.1, spheres and parallel-oriented cylinders tend to form terrace structures when the thickness of the block copolymer does not completely match the commensurability condition. Therefore, the thickness of the block copolymer for forming a homogeneous single layer is limited to a narrow window, which may be difficult to attain in largescale applications.

The orientations of microdomains in thin films are basically governed by the boundary condition between the block copolymer and the substrate or free surfaces, as described in the previous section. To obtain perpendicularly oriented microdomains, both substrate and free surfaces should be neutral to the polymer chains composing the block copolymer. In general, fulfillment of this condition cannot be attained spontaneously, and the following methods have been developed to orient the microdomains perpendicularly to the substrate. First, to control the interactions at the substrate interface, chemical modification of substrate surface is an effective way to achieve a neutral condition. Especially, polymer brush is widely applied. Brush polymers are a class of polymers that are adhered to a solid surface. By applying polymer brush composed of the same components as the block copolymer film, the surface becomes neutral for both of the blocks. For example, when using PS-*b*-PMMA block copolymer films, a monolayer of hydroxyl-terminated

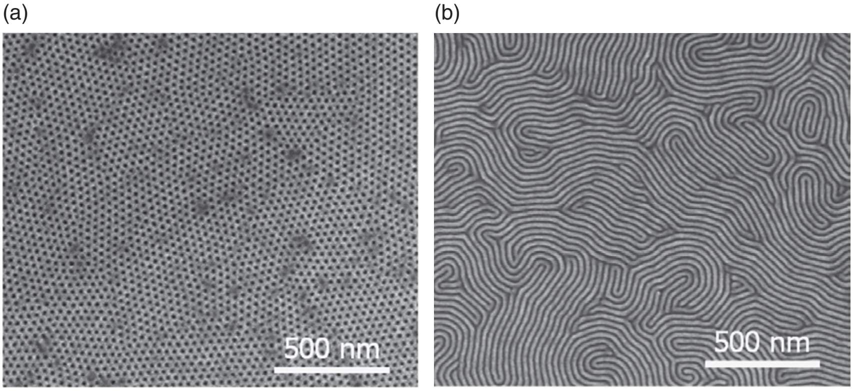
PS-*random*-PMMA copolymer is attached on the substrate surface (Mansky *et al.*, 1997a; Huang *et al.*, 1998; In 2006; Han *et al.,* 2008). To attain neutral condition, the fraction of PS and PMMA in the random copolymer should be tuned to match the composition of the block copolymer (In *et al.,* 2006; Han *et al.,* 2008).

On the other hand, attaining neutral condition for the interface at the free surface is rather difficult. In general, samples are annealed under vacuum or inert atmosphere to undergo microphase separation. Therefore, the parameter to control surface energy at the free surface is limited. PS and PMMA are a unique pair of polymers with very similar surface energy against air or vacuum, and can naturally satisfy a neutral condition at the free interface with thermal annealing (Mansky *et al.*, 1997b). This characteristic of PS and PMMA, in addition to the merit of etching selectivity as mentioned above, makes PS-*b*-PMMA one of the most appropriate materials for block copolymer lithography. Alternatively, annealing under a controlled solvent atmosphere can potentially provide a method to achieve a neutral condition at the free surface. In solvent annealing, the block copolymer thin film is exposed to solvent vapor to promote microphase separation. It has been reported that by choosing appropriate solvents and vapor pressures, an effective neutral condition can be achieved (Kim and Libera 1998; Kim *et al.,* 2004; Tokarev *et al.,* 2005).

8.3 Directed self-assembly of block copolymers

The spontaneous process of microphase separation of block copolymers leads to the formation of polycrystalline microdomain arrays with grains consisting of randomly oriented localized order, as demonstrated in Fig. 8.6. For this reason, the first demonstrations of block copolymer films in lithographic applications focused on the high pattern densities that could be attained, and ignored the issues of pattern registration and long-range order (Mansky *et al.*, 1995, 1996; Park *et al.,* 1997). However, to realize the full potential of self-assembled structures and to utilize them in more demanding applications such as magnetic storage or semiconductor devices, it is required to attain long-range ordering, good feature registration, and accurate placement with a very low number of defects. In the following section, we discuss methods to artificially direct the assembly of block copolymers to achieve long-range order in a predefined direction and to control feature placement in desired locations.

To control orientation and placement on block copolymer thin films, some type of guidance during the microdomain formation process is required. Several methods have been proposed for this purpose, including methods utilizing external fields such as share field (Angelescu *et al.*, 2004), electrical field (Morkved *et al.,* 1996) and methods utilizing lithographically predefined

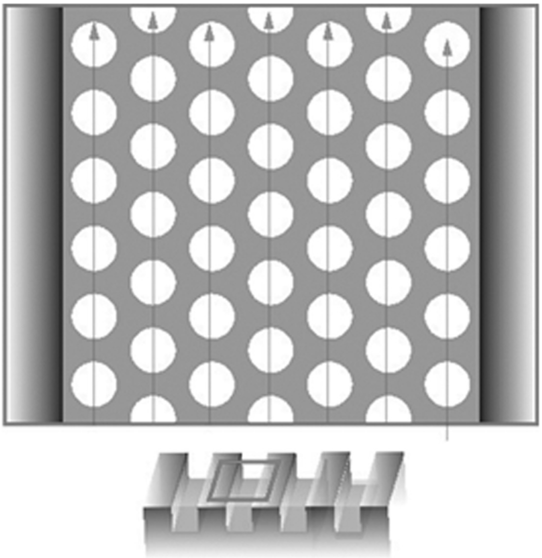


8.6 Top-view scanning electron micrographs of (a) cylindrical and (b) lamellar microdomain structures self-assembled by PS-*b*-PMMA. Scanning electron microscopy was carried out after etching PMMA phase by O_2 RIE. Without guidance, microdomains form poly-grain structure with short-range order.

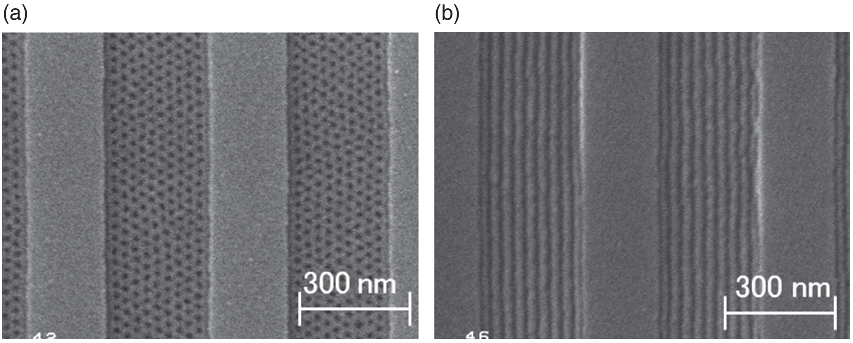
templates. The latter, which combines top-down conventional lithographic technology with bottom-up self-assembly, is a promising route to artificially and precisely direct self-assembly to achieve ordered patterns. In the following section, graphoepitaxy and chemical heteroepitaxy, which utilize topographically and chemically pre-patterned templates, respectively, are discussed as typical methods of directed self-assembly of block copolymer thin films.

Graphoepitaxy employs topographically patterned template to control orientation of crystal growth in thin films. The method was first reported to manipulate the orientation of inorganic crystals (Smith and Flanders 1978; Geis *et al.,* 1979). The application of graphoepitaxy to control ordering of block copolymers was demonstrated by Segalman *et al.* to align spherical microdomain formed by polystyrene-*block*-poly vinylpyridine (PS-*b*-PVP) (Segalman *et al.,* 2001).

As illustrated in Fig. 8.7, the graphoepitaxy of block copolymer typically employs templates with trenches. The trench width is generally designed in the range of a few times the natural domain spacing, d_0, of ordered block copolymers. The template is coated with a block copolymer film so that the block polymer material flows into the trench. As the block copolymer undergoes microphase separation by annealing, domains are aligned against the wall surface. Scanning electron microscope (SEM) photographs presented in Fig. 8.8a and 8.8b demonstrate the patterns obtained by self-assembling PS-*b*-PMMA to form a single layer of PMMA spheres and a layer of in-plane oriented PMMA cylinders in the grooves, respectively. Both results show that the block copolymer lattice aligns against the wall of the trench and forms a pattern with a single orientation in the confined space. A similar



8.7 Schematic illustration demonstrating concept of graphoepitaxiay. Block copolymer are self-assembled in trenches prepared by conventional lithography.



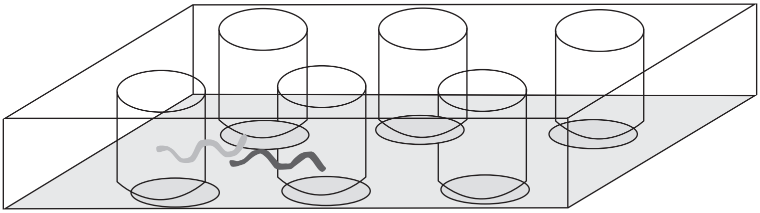
8.8 Top-view scanning electron micrographs of (a) spherical and (b) cylindrical microdomains self-assembled by PS-*b*-PMMA in trenches prepared on the Si substrate. Scanning electron microscopy was carried out after etching PMMA phase by O_2 RIE. Width and depth of the trench pattern: 275 and 50 nm, respectively.

effect was reported by employing perpendicularly oriented lamellae and cylindrical domains (Xiao *et al.,* 2005; Chen *et al.*, 2007; Park *et al.,* 2007b). The topological pattern can be also varied to include circular, rectangular, and triangular wells (Chai and Buriak, 2008), as well as cylindrical posts (Bita *et al.,* 2008).

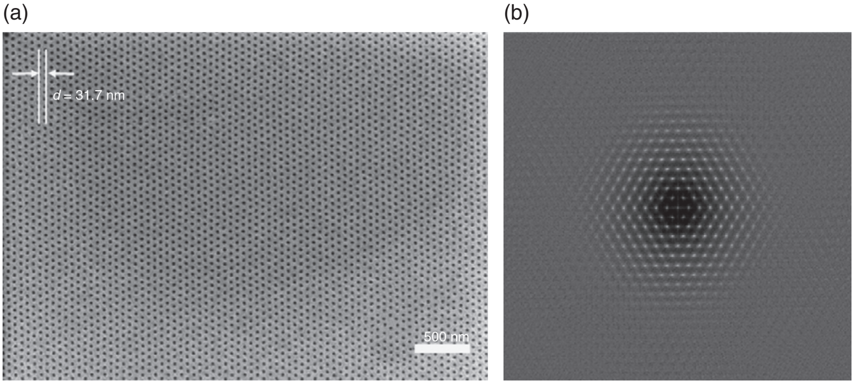
The effect of the wall on the arrangement of microdomains can be discussed primarily in terms of packing of domains in confined space with excluded-volume interaction. However, in block copolymer systems, preferential wetting of block copolymer chains to the wall surface has a significant effect. This effect becomes prominent when in-plain cylinders or perpendicularly oriented lamellae are assembled in the groove. When the wall surface is preferentially wet by one of the components of the block copolymer, the cylinders/lamellae are aligned parallel to the wall surface. On the other hand, when the wall surface is neutral to both of the components, cylinders/lamellae are reported to orient perpendicularly to the wall (Xiao *et al.,* 2005; Park *et al.,* 2007b).

Chemical heteroepitaxy, or chemical registration, of block copolymer thin films utilizes chemically pre-patterned templates with surface energy contrast that directs the registration of the block copolymer microdomains. The effect of a chemically patterned surface on the orientation of block copolymer microdomains was first reported by Rockford *et al.* (1999). The surface prepared by Rockford consisted of alternating stripes of silicon oxide and Au. The chemical difference in surface pattern directed the phase separated structure of PS-*b*-PMMA. Then, Nealey *et al.* have successfully directed lamellar- (Kim *et al.,* 2003) and cylinder-forming (Edwards *et al.,* 2006; Park *et al.,* 2007a) block copolymer self-assembly in registration with the lithographic pattern defined in the chemically modified surfaces with a one-to-one correspondence.

Figure 8.9 schematically illustrates the concept of the technique considering PS-*b*-PMMA with perpendicularly oriented PMMA cylindrical microdomain as an example. The chemical contrast template is prepared by patterning the surface of a substrate in the circular regions with a hexagonal closed pack (hcp) arrangement. The surface of these circular regions are designed to have higher affinity, i.e. lower surface energy to PMMA, while the surrounding matrix is tuned for preferential wetting by PS. The lattice constant of the hcp lattice of the template, d_{sub}, is designed to be the same as, or to be a multiple of, that of the natural spacing d_0 of PS-*b*-PMMA. When PS-*b*-PMMA is applied on the template and microphase separates to form PMMA cylinders with an



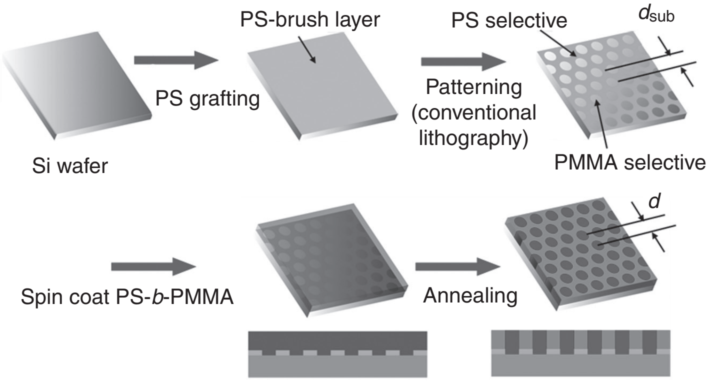
8.9 Schematic illustration demonstrating concept of chemical heteroepitaxy of block copolymer self-assembly.



8.10 (a) A top-view scanning electron micrograph of cylindrical microdomain structure of PS-*b*-PMMA with d_0 = 32 nm self-assembled perpendicularly on a chemically patterned substrate with d_{sub} = 32 nm with 1:1 correspondence, and (b) a corresponding 2D-FFT image. The weight averaged molecular weight of PS-*b*-PMMA and weight fraction of its PS block were 6.7×10^4 and 0.69, respectively.

hcp lattice, the PMMA cylinders are registered on the circular regions. The registration is promoted by minimizing the free energy of the film by placing the PMMA cylinder on the circular regions and PS matrix on the surrounding area. Figure 8.10a demonstrates the effect of chemical heteroepitaxy. This SEM shows a top-down view of PS-*b*-PMMA thin film self-assembled on a chemical contrast template to form perpendicular cylinders with a lattice constant d = 32 nm (full pitch). Underlying substrate is chemically patterned with $d_{\text{sub}} = d_0$ = 32 nm. The PMMA cylindrical microdomains are perfectly arranged in hcp lattice, which can also be confirmed by Fast Fourier Tranform (FFT) power spectrum of the SEM (Fig. 8.10b).

A typical procedure to prepare the chemical template for PS-*b*-PMMA is described in Fig. 8.11. First, PS is adhered to a Si substrate to form a thin layer of PS, i.e. PS-brush. This can be conveniently processed by spin coating a hydroxyl-terminated PS (PS-OH) on the fresh surface of the oxidized Si. The PS film is then thermally annealed to bond the PS-OH to the Si surface. Lastly, the excess of non-bonded PS-OH is removed by rinsing the substrate with a suitable solvent. The next step is to pattern the PS-brush monolayer. Conventional top-down type lithography can be applied for this purpose, including photolithography and e-beam lithography. A resist material is coated on the PS-brush surface, exposed, and developed to form circular holes in the resist film, which can be used as a mask for a subsequent O_2 RIE of PS-brush. The template is ready after stripping the resist material. The template prepared by this method has chemical pattern composed of circular regions with oxidized Si or oxidized PS-brush surface surrounded



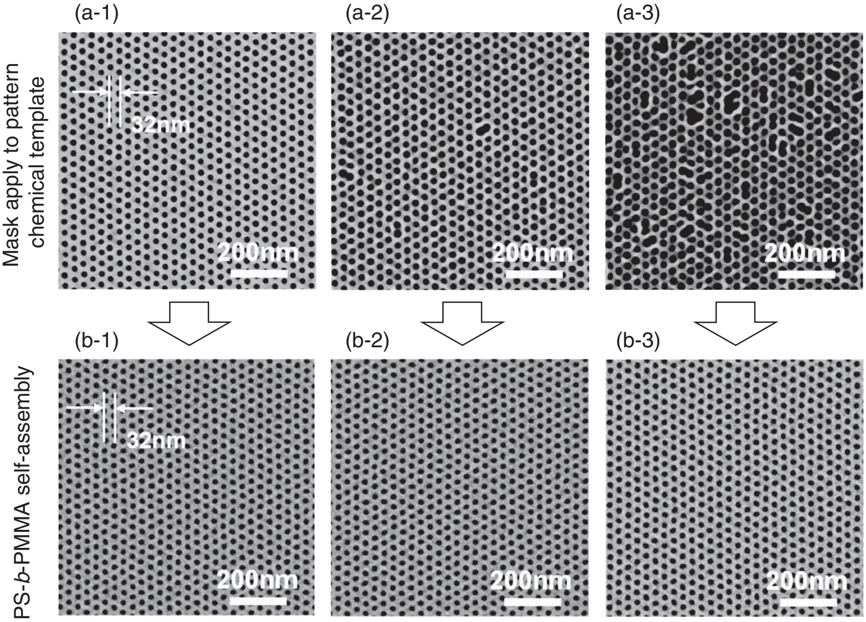
8.11 Schematically illustrated chemical heteroepitaxy process of PS-*b*-PMMA with chemically pre-patterned template utilizing PS-brush surface.

by PS-brush surface. The former regions, which have higher hydrophobicity than the surrounding, are expected to have higher affinity to PMMA, while surrounding PS-brush surface has higher affinity to PS.

Surface treatment/modification has to be tuned so that the chemical contrast of the template matches the components of block copolymer. Besides polymer brushes, a wide variety of surface modification techniques and/or substrates can be employed for preparing the chemical template, including self-assembled monolayers (Kim *et al.,* 2003), cross-linkable polymers (Liu *et al.,* 2011) and organosilicate layers cured at controlled temperatures (Suh *et al.,* 2010).

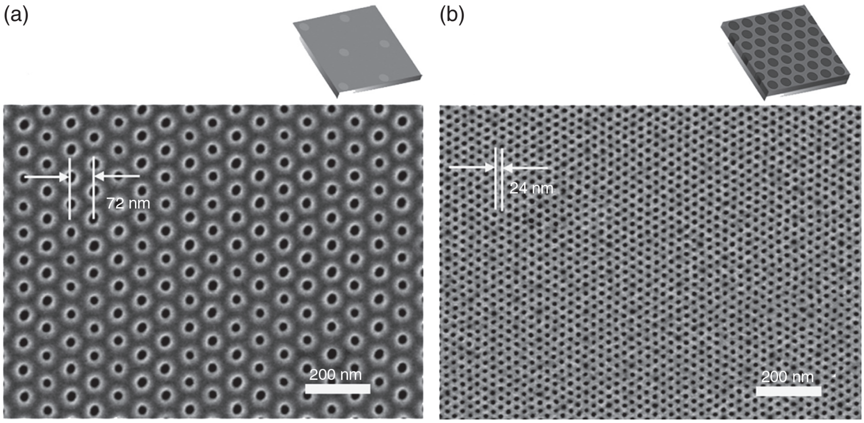
As demonstrated in Fig. 8.10, single grain structures with long-range order and fewer defects can be attained by applying chemical registration. Moreover, the method provides two characteristics especially appealing for lithographic applications: pattern rectification, i.e. pattern clean-up or pattern healing, and pattern feature density multiplication.

The structure and dimensions of self-assembled block copolymer domains are primarily determined by the characteristics of block copolymer molecules including composition, molecular weight and segregation strength. Block copolymers with sufficiently small molecular weight distribution form domains with homogeneous dimension and flat interface. These characteristics can be applied to rectify the short-range errors in pattern prepared by conventional lithography. An example of the pattern rectification effect is demonstrated in Fig. 8.12. Top-down SEMs presented on the top row (Fig. 8.12 (a-1), (a-2) and (a-3)) present the resist patterns prepared by e-beam lithography that have been employed to pattern chemical templates. The SEMs on the bottom low (Fig. 8.12 (b-1), (b-2) and (b-3)) demonstrate the patterns obtained by self-assembling PS-*b*-PMMA to form PMMA cylinders on the



8.12 (a-1) to (a-3) Top-view scanning electron micrographs of EB resist with different degree of defects contents, which were employed to pattern PS graft layer on Si wafer surface for chemically pre-patterned templates with d_{sub} = 32 nm. (b-1) to (b-3) Top-view scanning electron micrographs of cylinder structures of PS-*b*-PMMA with d_0 = 32 nm self-assembled perpendicularly on the chemically pre-patterned templates prepared by applying EB resist patterns presented in (a-1) to (a-3), respectively. The weight averaged molecular weight of PS-*b*-PMMA and weight fraction of its PS block were 6.7×10^4 and 0.69, respectively.

chemical templates in perpendicular orientation. In this particular experiment, the resist patterns were artificially designed so that the pattern quality changes in short-range. In spite of the existence of significant defects in some of the resist patterns, the pattern obtained by self-assembling PS-*b*-PMMA showed a more uniform pattern quality. A comparison of Fig. 8.12 (a-3) and (b-3) reveals that variations in pattern position and point defects in the resist pattern can be corrected by the block copolymer self-assembly. This pattern rectification effect has also been reported to be effective at reducing line-edge roughness in resist patterns (Daoulas *et al.,* 2008; Stoykovich *et al.,* 2010) and at correcting width variations of line-and-space patterns (Edwards *et al.,* 2007). As critical dimensions reach smaller dimensions, conventional lithographic techniques are approaching their resolution limits. In such cases, the pattern may contain a significant number of errors and variations. The pattern rectification aspect of DSA can be regarded as an effective way to correct the pattern in such cases.



8.13 (a) A top-view scanning electron micrograph of EB resist employed to pattern PS graft layer on Si wafer surface for a chemically pre-patterned template with d_{sub} = 72 nm. (b) A top-view scanning electron micrograph of cylinder structures of PS-*b*-PMMA with d_0 = 24 nm self-assembled perpendicularly on the chemically pre-patterned template. The weight averaged molecular weight of PS-*b*-PMMA and weight fraction of its PS block were 4.8×10^4 and 0.75, respectively.

It is also possible to extend the use of chemical contrast patterns to achieve feature density multiplication or, in another word, frequency multiplication of sparse chemical template patterns (Ruiz *et al.,* 2008; Tada *et al.,* 2008). Figure 8.13 demonstrates a nine-fold pattern density multiplication of a hexagonal lattice pattern. Figure 8.13a shows a SEM of the EB resist pattern employed to prepare the chemical template with d_{sub} = 72 nm. Figure 8.13b shows a SEM of a thin film of PS-*b*-PMMA with the natural domain spacing of d_0 = 24 nm, self-assembled on the pre-patterned chemical template. The self-assembled film is well-aligned to the template pattern with long-range order. The lattice spacing of the self-assembled PMMA cylindrical microdomain was measured to be d = 24 nm, which was a third of the value of the lattice spacing, d_{sub} = 72 nm, of the chemically pre-patterned substrate. The block copolymer features interpolated the chemical contrast features to achieve a nine-fold density multiplication. Needless to point out, feature density multiplication can be a promising route to extend the limitations in current lithographic techniques to reach smaller dimensions.

In the above section, chemical heteroepitaxy of block copolymer was discussed considering mainly hcp patterns with perpendicularly oriented cylindrical domains. However, chemical heteroepitaxy can be applied to form hcp pattern by arranging sphere-forming materials (Park *et al.,* 2008b; Xiao *et al.,* 2009; Tada *et al.,* 2012). Furthermore, chemical heteroepitaxy can also be used to form line-and-space patterns employing perpendicularly oriented lamellae (Kim *et al.,* 2003; Edwards *et al.,* 2007; Cheng *et al.,* 2008; Stoykovich *et al.,* 2010; Liu *et al.,* 2011) or in-plane oriented cylindrical domains (Edwards *et al.,* 2006). Additionally, it has been reported that changes in morphology can be induced by chemical heteroepitaxy (La *et al.*, 2005). Moreover, chemical heteroepitaxy has also been successfully applied to complex patterns, such as bended lines, isolated lines, and junctions, which basically demonstrates that the method can be applicable in semiconductor logic patterning (Stoykovich *et al.,* 2007; Kang *et al.,* 2008; Liu *et al.,* 2010; Ruiz *et al.,* 2011).

8.4 Programmable three-dimensional lithography

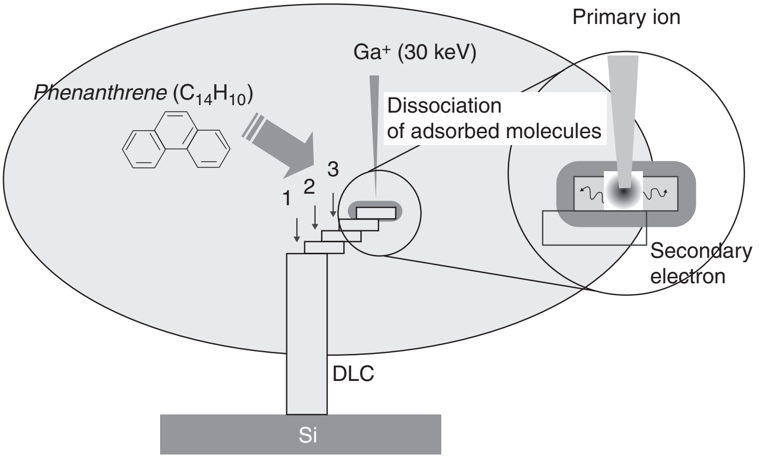
Two-dimensional nanolithography using EB and FIB is described in Chapters 3 and 4. In this section we will describe three-dimensional (3D) fabrication by FIB chemical-vapor-deposition (FIB-CVD) using a 3D pattern degeneration system. Next, 3D nanostructure fabrication will be described by EB-CVD using a SEM with a beam control system.

8.4.1 Three-dimensional lithography using FIB-CVD

Matsui *et al*. (2000) demonstrated 3D nanostructure fabrication by FIB-CVD. In this experiment, two commercially available FIB systems (SMI9200, SMI2050, SII Nanotechnology Inc., Tokyo, Japan) with a Ga^+ ion beam operating at 30 keV were used. The FIB-CVD used a precursor of *phenanthrene* ($C_{14}H_{10}$) as the carbon source material. The beam diameter of SMI9200 was about 7 nm and that of SMI2050 was about 5 nm. The SMI9200 system was equipped with two gas sources in order to increase the gas pressure. The top of the gas nozzles faced each other and was directed at the beam point. The nozzles faced each other and were directed at the beam point. The nozzles were set a distance of 40 μm from each other and positioned about 300 μm above the substrate surface. The inside diameter of a nozzle was 0.3 mm. The *phenanthrene* gas pressure during pillar growth was typically 5×10^{-5} Pa in the specimen chamber, but the local gas pressure at the beam point was expected to be much higher. The crucible of the source was heated to 85°C. The SMI2050 system, on the other hand, was equipped with a single gas nozzle. The FIB was scanned in order to be able to write the desired pattern via computer control, and the ion dose was adjusted to deposit a film of the desired thickness. The experiments were carried out at room temperature on a silicon substrate.

The deposited film was characterized by observing it with a transmission electron microscope (TEM). A thin film of carbon (200 nm thick) was deposited on a silicon substrate by 30 keV Ga^+ FIB using *phenanthrene* precursor gas. The cross-sections of the structures created and their electron diffraction patterns were observed by using a 300 kV TEM. There were no crystal structures in the TEM images and diffraction patterns. It was therefore concluded that the deposited film was amorphous carbon (a-C). Raman spectra of the a-C films were measured at room temperature with 514.5 nm line of an argon ion laser. The Raman spectra were recorded using a monochromator equipped with a Charge Coupled Device (CCD) multi-channel detector. Raman spectra were measured at 0.1–1.0 mW to avoid thermal decomposition of the samples. A relatively sharp Raman band at 1550 cm^{-1} and a broad-shouldered band at around 1400 cm^{-1} were observed in the spectra excited by the 514.5 nm line. Two Raman bands were plotted after Gaussian line shape analysis. These Raman bands, located at 1550 cm^{-1} and 1400 cm^{-1}, originate from the trigonal (sp^2) bonding structure of graphite and tetrahedral (sp^3) bonding structure of diamond. This result suggests that the a-C film deposited by FIB-CVD is diamond-like amorphous carbon (DLC), which has attracted attention due to its hardness, chemical inertness, and optical transparency. The coordination of carbon atoms in the carbon-based material formed by CVD of *phenanthrene* assisted by Ga FIB was investigated by the measurement of near-edge X-ray absorption fine-structure spectra of the carbon *K*-edge over the excitation energy range 275–320 eV (Kanda *et al.,* 2006). The novel peak observed at 289.0 eV was assigned to the $1s \rightarrow \sigma^*$ transition of carbon neighboring to the residue gallium. The material formed by this method was found to be Ga-doped diamond-like-carbon, which consists of a high sp^3 hybridized carbon. The atomic fraction of the FIB-CVD DLC film has been determined as C: Ga: H = 87.4 at%: 3.6 at%: 9.0 at% using Rutherford backscattering spectrometry (RBS) and elastic recoil detection analysis (ERDA) (Igaki *et al.*, 2007). Hydrogen content of FIB-CVD DLC film was relatively lower than that of DLC films formed by other CVD methods.

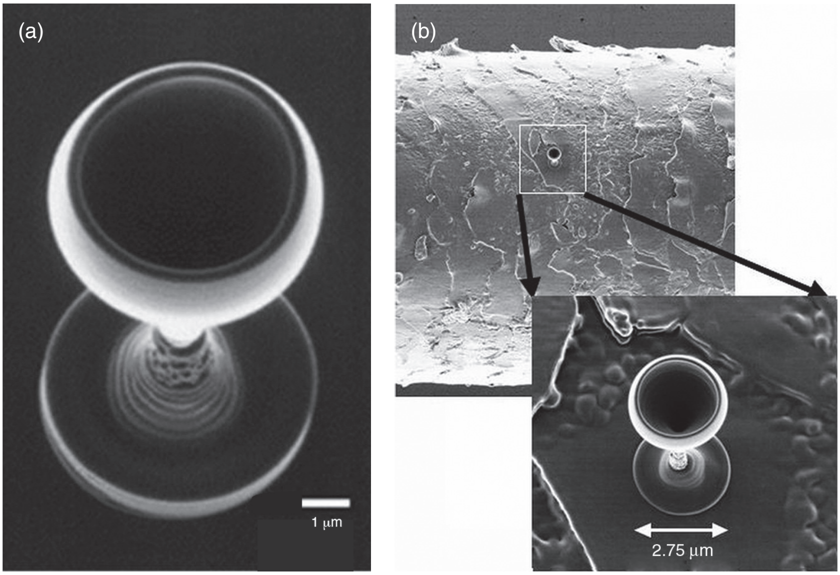
The key issue in realizing such 3-D nanostructures is the short penetration depth of the ions (a few tens of nm) into the target material, where the penetration depth of the ions is much shorter than that of electrons (several hundreds of nm). This short penetration depth reduces the dispersion area of the secondary electrons, and so the deposition area is restricted to roughly several tens nanometers. A 3-D structure usually contains overhang structures and hollows. Gradual position scanning of the ion beam during the CVD process causes the position of the growth region around the beam point to shift. When the beam point reaches the edge of the wall, secondary electrons appear at the side of the wall and just below the top surface. The DLC then starts to grow laterally; the width of the vertical growth is also about 80 nm. Therefore, combining the lateral growth mode with rotating beam scanning, it is possible to obtain 3-D structures with rotational symmetry like a wine


Primary ion
Ga+ (30 keV)
Phenanthrene ($C_{14}H_{10}$)
Dissociation of adsorbed molecules
1
2
3
Secondary electron
DLC
Si

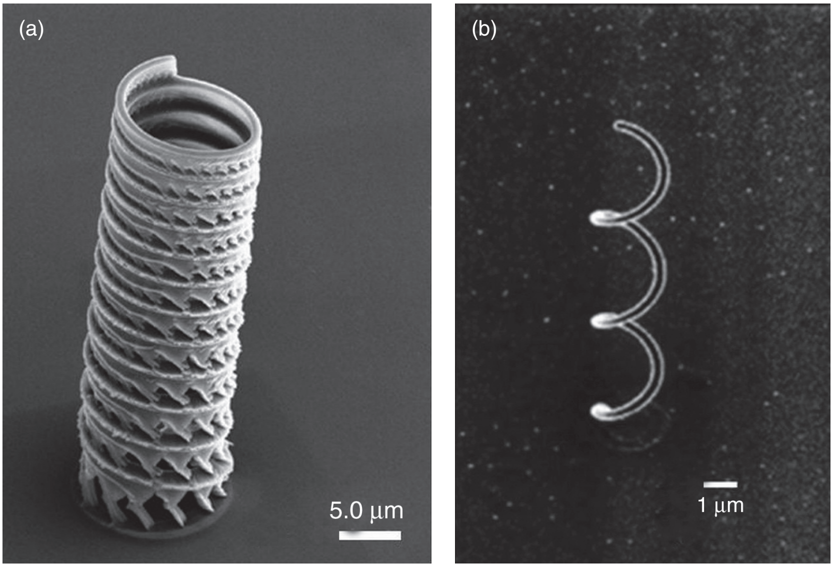
8.14 Fabrication process for three-dimensional nanostructure by FIB-CVD.

glass. The process of fabricating three-dimensional structures by FIB-CVD is illustrated in Fig. 8.14 (Matsui *et al.,* 2000). In FIB-CVD processes, the beam is scanned in digital mode. First, a pillar is formed on the substrate by fixing the beam position (Position 1). After that, the beam position is moved to within a diameter of the pillar (Position 2) and then fixed until the deposited terrace thickness exceeds the range of the ions (a few tens of nm). This process is repeated to make three-dimensional structures. The key point to making three-dimensional structures is to adjust the beam-scan-speed so that the ion beam remains within the deposited terrace, which means that the terrace thickness always exceeds the range of the ions. The growth in the x and y- directions is controlled by both beam-deflectors. The growth in the z-direction is determined by the deposition rate; that is, the height of the structure is proportional to an irradiation-time when a deposition rate is constant.

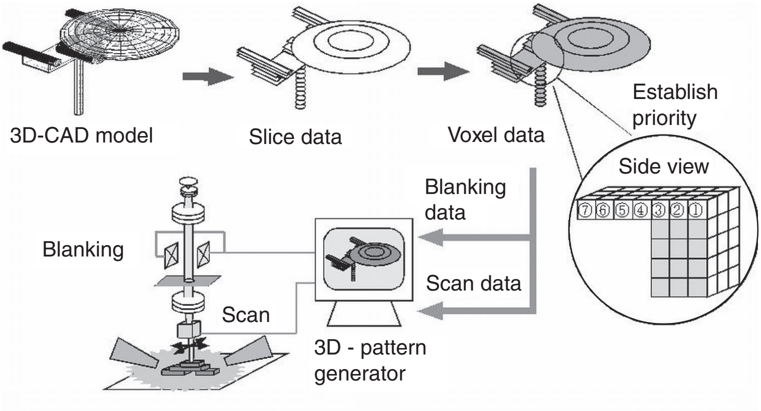
To demonstrate the possibilities of this field, a 'micro-wine-glass' was created on a Si substrate and a human hair as a work of microstructure plastic arts, as shown in Fig. 8.15a and 8.15b. A micro-wine glass with an external diameter 2.75 μm and a height of 12 μm was formed. The fabrication time was 600 s at a beam current 16 pA. This beautiful micro-wine glass shows the potential of the field of microstructure plastic art. A 'micro leaning tower of Pisa' was also fabricated on an Si substrate, as shown in Fig. 8.16a. Figure 8.16b shows a microcoil with a coil diameter of 0.6 μm, a coil pitch of 0.7 μm coil pitch, and a linewidth 0.08 μm. The exposure time was 40 s at a beam current 0.4 pA. The results show that FIB-CVD is a highly promising technique for realizing parts of a microsystem.



8.15 Micro-wine glass with an external diameter of 2.75 μm and a height of 12 μm on (a) a Si substrate and (b) a human hair.



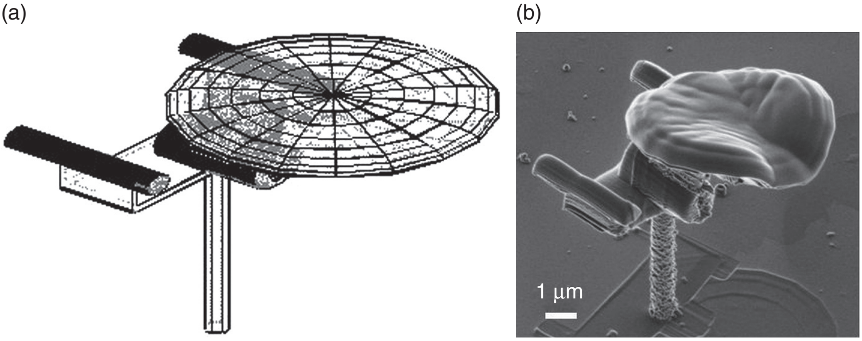
8.16 (a) 'Micro Leaning Tower of Pisa.' (b) Microcoil with a coil diameter of 0.6 μm, a coil pitch of 0.7 μm, and a linewidth of 0.08 μm.



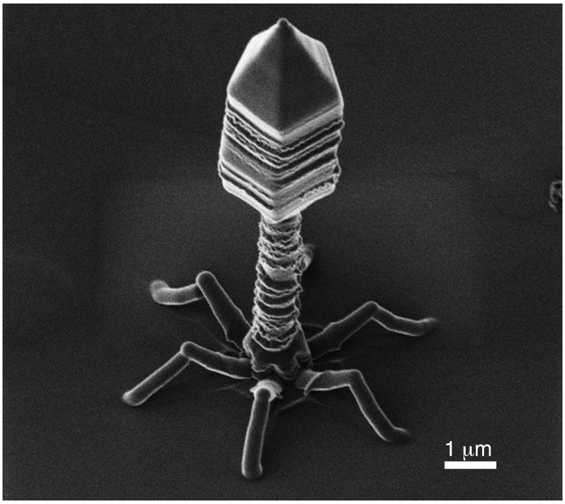
8.17 Data flow of 3-D pattern-generating system for FIB-CVD.

The 3-D structure is builtup as a multilayer structure. In the first step of this 3-D pattern-generating system, a 3-D model of the structure, designed using a 3-D Computer Aided Design (CAD) system (3-D DXF format), is needed. In this case, a structure shaped like a pendulum was realized. The 3-D CAD model, which is a surface model, is cut into several slices, as shown in Fig. 8.17. The thickness of the slices depends upon the resolution in the z- direction (the vertical direction). The x and y coordinates of the slices are then used to create the scan data (voxel data). To fabricate the overhanging structure, the ion beam must irradiate the correct positions in the correct order. If the ion beam irradiates a voxel located in mid-air without a support layer, the ions intended for the voxel will be deposited on the substrate. Therefore, the sequence of irradiation is determined, as shown in Fig. 8.17. The scan data and blanking signal therefore include the scan sequence, the dwell time, the interval time, and the irradiation pitch. These parameters are calculated from the beam diameter, x-y resolution, and z resolution of fabrication. The z resolution is proportional to the dwell time and inversely proportional to the square of the irradiation pitch. The scan data are passed to the beam-deflector of the FIB-CVD as the blanking data. The blanking signal controls the dwell time and interval time of the ion beam.

Figure 8.18 shows a 3-D CAD model and a scanning ion microscope (SIM) image of the Starship Enterprise NCC-1701D (from the television series *Star Trek*), which was fabricated by FIB-CVD at 10~20 pA (Hoshino *et al.,* 2003). The nano-spaceship is 8.8 μm long and was realized at about a 1:100 000 000 scale on silicon substrate. The dwell time (t_d), interval time (t_i), irradiation pitch (p), and total process time (t_p), were 80 μs, 150 μs, 2.4 nm, and 2.5 h, respectively. The horizontal overhang structure was fabricated



8.18 'Micro Starship Enterprise NCC-1701D,' 8.8 μm long. (a) 3-D CAD model, (b) SIM image (tilt 45°).



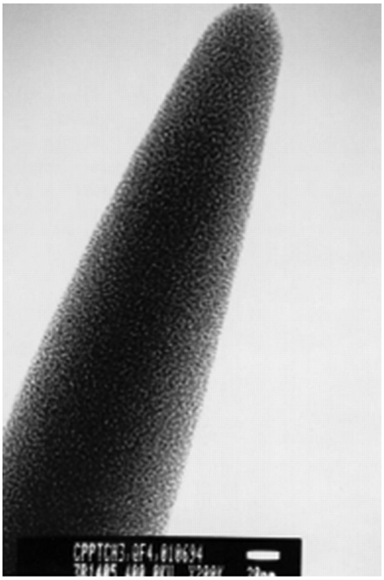
8.19 T-4 bacteriophage.

successfully. Figure 8.19 shows a 'nano-T-4 bacteriophage,' which is an artificial version of the virus fabricated by FIB-CVD on silicon surface. The size of the artificial nano 'T-4 bacteriophage' is about ten times as that of the real virus.

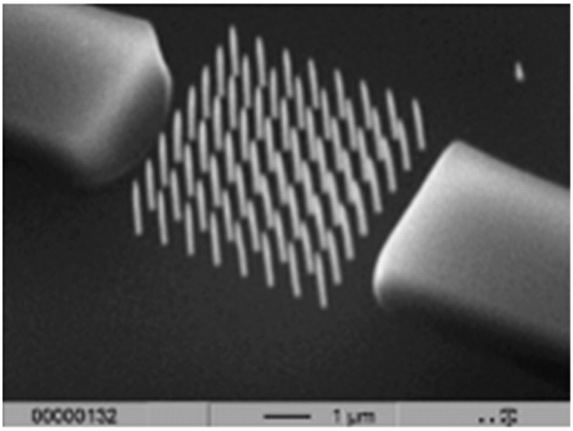
8.4.2 Three-dimensional lithography using EB-CVD

Broers *et al.* (1976) were able to deposit nanometer scale carbon contamination structures in a scanning TEM with electrons of 45 keV energy. Matsui and Mori (1984) obtained the first metal-containing deposits by introducing a metal–organic gas ($Cr(C_6H_6)_2$) vapor in the vacuum chamber of a scanning electron microscope. Since the first experiments, many EB-CVD metal-containing structures have been deposited under a wide range of experimental conditions during the past 30 years (Matsui and Mori 1986; Matsui and Ichihasi 1988; Matsui *et al.,* 1989; Koops *et al.,* 1988, 2001a, 2001b). The first metal-containing three-dimensional nanostructures were deposited by Koops *et al*. (1998), using the focused EB of a conventional SEM.

Koops *et al*. (2001b) reported the experiments using three-dimensional EB-CVD. This technique is very well suited for rapid prototyping of photonic crystals. It is used to grow arrays of rods from nanocrystalline material having a very high aspect ratio over 15 under computer control. The nanolithography tool in use is based on a scanning electron microscope Jeol JSM 6300 F with beam control system. All possible precautions have been taken to compensate for thermal drifts, mechanical vibrations, and magnetic interference by external compensating systems. These measures allow the high precision fabrication of two-dimensional arrays of dielectric rods with the dimensions required for photonic crystals. For the precursor material supply to the waveguide sample, a specially designed environmental chamber is used (Koops *et al.,* 2001a). The EB-CVD process renders 2 nm edge roughness of the deposited rods. This is a surface quality unreachable by standard lithography or beam-assisted etching techniques. It allows for minimizing



8.20 The high resolution TEM image proves the wall roughness being < 2 nm of the deposited rod if a platinum containing precursor material is used.

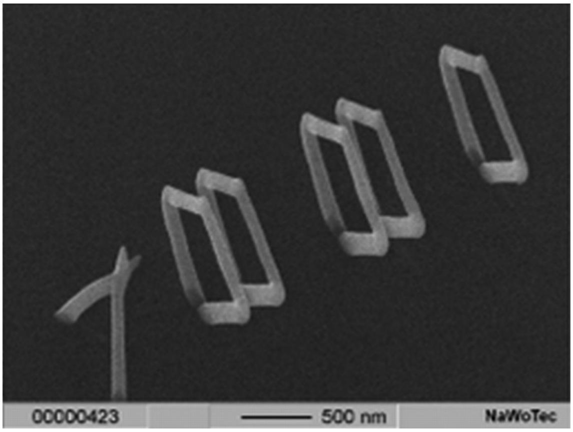


8.21 A cubic photonic crystal is constructed in a prefabricated gap in a PMMA waveguide for investigation of the transmission properties of the phonic crystal.

the scattering of photons. Figure 8.20 proves with a high resolution TEM image the wall roughness being < 2 nm of the rod grown by EB-CVD using cyclopentadienyl-platinum-trimethyl as a precursor material. EB-CVD constructs the rods from nanocrystals of 2 nm diameter embedded into a carbonaceous matrix. This results in an edge roughness of below 2 nm, which is an outstanding feature of this process.

Photonic crystals with cubic, hexagonal, or honeycomb elementary cells are fabricated under computer control. Using 25 keV electrons and 20 pA current, a crystal of 100 rods is grown in 40 min. For this time, the sample in the microscope is not allowed to drift more than 20 nm. Figure 8.21 shows a cubic photonic crystal, which is constructed into a prefabricated gap in a PMMA waveguide for investigation of the transmission properties of the photonic crystal. The pitch is 580 μm, the rod diameter is 150nm, and the height is 2.5 μm. Characterization of photonic crystals and related devices is a very challenging task. This is due to the small size of the devices, having diameters of 3–6 λ, which is 4.5–9 μm at $\lambda = 1.5$ μm. Monomode fibers have a core diameter of 9 μm. Therefore, a well-designed optical beam converter is required to couple light into photonic crystals. To measure the polarization-dependent transmission of photonic crystals, a sophisticated metrology set-up was used. The transmission measurement did prove a band gap filter in the regime from1250 to 1650 nm.

EB-CVD is a rapid prototyping technology with a very high resolution and accuracy to investigate the activity and characteristics of photonic crystals and devices. The initial promising results prove the applicability of the technology and the optical action of the devices. EB-CVD can build similar



8.22 Smith-Purcell electron optics realized with EB-CVD.

structures in the sub-μm dimensions with wires of 80–200 nm. Figure 8.22 shows a three-dimensional build-up with conducting material resembling a field emission electron source, having an emitter tip and an extractor ring around it (Floreani *et al.*, 2001). The structure has a capacitance of 24 aF between the tip and the ring.

8.5 Conclusion

In this chapter, we reviewed directed assembly lithography using self-assembly of block copolymer and programmable 3D lithography using FIB- and EB-CVD. In directed self-assembly of block copolymer, we have discussed microphase separation of block copolymers in the bulk and in thin film, block copolymer lithography, and directed self-assembly of block copolymer utilizing lithographically predefined templates, which combined top-down conventional lithographic technology with bottom-up self-assembly. In programmable three-dimensional lithography using FIB- and EB-CVD, 3D nanostructure fabrication has been demonstrated by 30 keV Ga^+ FIB-CVD and a *phenanthrene* ($C_{14}H_{10}$) source as a precursor. The film deposited on a silicon substrate was characterized using a transmission microscope and Raman spectra. This characterization indicated that the deposited film is DLC, which has attracted attention due to its hardness, chemical inertness, and optical transparency. As an application of three dimension by EB-CVD using a 10 keV SEM, nano-optics with the rapid prototyping of photonic crystals and nano-electronics with micro-triodes have been described. These results demonstrate that FIB- and EB-CVD are one of the key technologies needed to make 3-D nanodevices that can be used in the field of electronics, mechanics, optics and biology.

8.6 References

Angelescu D E, Waller J H, Adamson D H, Deshpande P, Chou S Y, Register R A and Chaikin P M (2004), 'Macroscopic orientation of block copolymer cylinders in single-layer films by shearing', *Adv. Mater*., **16**, 1736–1740.

Bita I, Yang J K W, Jung Y S, Ross C A, Thomas E L and Berggren K K (2008), 'Graphoepitaxy of self-assembled block copolymers on two-dimensional periodic patterned templates', *Science*, **321**.

Broers A N, Molzen W W, Cuomo J J and Wittels N D (1976), 'Electron-beam fabrication of 80-Å metal structure', *Appl. Phys. Lett*., **29**, 596–599.

Chai J and Buriak J M (2008), 'Using cylindrical domains of block copolymers to self-assemble and align metallic nanowires', *ACS Nano*, **2**, 489.

Chen F, Akasaka S, Inoue T, Takenaka M, Hasegawa H and Yoshida H (2007), 'Ordering cylindrical microdomains for binary blends of block copolymers with graphoepitaxy', *Macromol. Rapid. Commun*., **28**, 2137–2144.

Cheng J Y, Rettner C T, Sanders D P, Kim H -C, Hinsberg W D (2008), 'Dense self-assembly on sparse chemical patterns: rectifying and multiplying lithographic patterns using block copolymers', *Adv. Mater.*, **20**, 3155–3158.

Daoulas K C, Müller M, Stoykovich M P, Kang H, de Pablo J J and Nealey P F (2008), 'Directed copolymer assembly on chemical substrate patterns: A phenomenological and single-chain-in-mean-field simulations study of the influence of roughness in the substrate pattern', *Langmuir*, **24**, 1284–1295.

Edwards E W, Stoykovich M P, Solak H H and Nealey P F (2006), 'Long-range order and orientation of cylinder-forming block copolymers on chemically nanopatterned striped surfaces', *Macromolecules*, **39**, 3598–3607.

Edwards E W, Muller M, Stoykovich M P, Solak H H, de Pablo J J and Nealey P F (2007), 'Dimensions and shapes of block copolymer domains assembled on lithographically defined chemically patterned substrates', *Macromolecules*, **40**, 90–96.

Fasolka M J and Mayes M A (2001), 'Block copolymer thin films: physics and applications', *Annu. Rev. Mater. Res*., **3**, 323–355.

Floreani F, Koops H W and Elsäβer (2001), 'Operation of high power field emitters fabricated with electron beam deposition and concept of a miniaturized free electron laser', *Microelectron. Eng*., **57–58**, 1009–1016.

Geis M W, Flanders D C and Smith H I (1979), 'Graphoepitaxy of silicon on fused silica using surface micropatterns and laser crystallization', *J. Vac. Sci Technol*., **16**, 1640–1643.

Hamley I W (1998), *The Physics of Block Copolymers*, Oxford University Press, New York.

Hamley I W (2009), 'Ordering in thin films of block copolymers: fundamentals to potential applications', *Prog. Polym. Sci.*, **34**, 1161–1210.

Han E, Stuen K O, La Y-H, Nealey P F and Gopalan P (2008), 'Effect of composition of substrate-modifying random copolymers on the orientation of symmetric and asymmetric diblock copolymer domains', *Macromolecules*, **41**, 9090–9097.

Harrison C, Park M, Chaikin P, Register R A, Adamson D H and Yao N (1998), 'Depth profiling block copolymer microdomains', *Macromolecules*, **31**, 2185–2189.

Harrison C, Chaikin P M, Huse D A, Register R A and Adamson D H (2000), 'Reducing substrate pinning of block copolymer microdomains with a buffer layer of polymer brushes', *Macromolecules*, **33**, 857–865.

Hoshino T, Watanabe K, Kometani R, Morita T, Kanda K, Haruyama Y, Kaito T, Fujita J, Ishida M, Ochiai Y and Matsui S (2003), 'Development of three-dimensional pattern-generating system for focused-ion-beam chemical-vapor deposition', *J. Vac. Sci. Technol. B,* **21**, 2732–2736.

Huang E, Russell T P, Harrison C, Chaikin P M, Register R A, Hawker C J and Mays J (1998), 'Using surface active random copolymers to control the domain orientation in diblock copolymer thin films', *Macromolecules*, **31**, 7641–7650.

Igaki J, Saikubo A, Kometani R, Kanada K, Suzuki T, Niihara K and Matsui S (2007), 'Elementary analysis of diamond-like carbon film formed by focused-ion-beam chemical vapor deposition', *Jpn. J. Appl. Phys*., **46**, 8003–8004.

In I, La Y H, Park S M, Nealey P F and Gopalan P (2006), 'Side-chain-grafted random copolymer brushes as neutral surfaces for controlling the orientation of block copolymer microdomains in thin films', *Langmuir*, **22**, 7855–7860.

Kang H, Crag G S W and Nealey P F (2008), 'Directed assembly of asymmetric ternary block copolymer-homopolymer blends using symmetric block copolymer into checkerboard trimming chemical pattern', *J. Vac. Sci. Technol. B*, **26**, 2495–2499.

Kanda K, Igaki J, Kato Y, Kometani R, Saikubo A and Matsui S (2006), 'NEXAFS study on carbon-based material formed by focused-ion-beam chemical-vapor-deposition', *Radiat. Phys. Chem.*, **75**, 1850–1854.

Khandpur A K, Forster S, Bates F S, Hamley I W, Ryan A J, Bras W, Almdal K and Mortensen K (1995), 'Polyisoprene-polystyrene diblock copolymer phase diagram near the order-disorder transition', *Macromolecules*, **28**, 8796–8806.

Kim G and Libera M (1998), 'Kinetic constraints on the development of surface microstructure in sbs thin films', *Macromolecules*, **31**, 2670–2672.

Kim S H, Misner M J, Xu T, Kimura M and Russell T P (2004), 'Highly oriented and ordered arrays from block copolymers via solvent evaporation', *Adv. Mater.*, **16**, 226–231.

Kim S O, Solak H H, Stoykovich M P, Ferrier N J, de Pablo J J and Nealey P F (2003), 'Epitaxial self-assembly of block copolymers on lithographically defined nanopatterned substrates', *Nature*, **424**, 411–414.

Koops H W P, Weiel R, Kern D P and Baum T H (1988), 'High-resolution electron-beam induced deposition', *J. Vac. Sci. Technol. B*, **6**, 477–481.

Koops H W, P Reinhardt A, Klabunde F, Kaya A and Plontke R (2001a), 'Vapor supply manifold for additive nanolithography with electron beam induced deposition', *Microelectron. Eng*., **57–58**, 909–913.

Koops H W P, Hoinks O E, Honsberg M E W, Schmidt R, Blum R, Böttger G, Kuligk A and Eich M (2001b), 'Two-dimensional photonic crystals produced by additive nanolithography with electron beam-induced deposition act as filters in the infrared', *Microelectron. Eng*., **57–58**, 995–1001.

La Y-H, Edwards E W, Park S-M and Nealey P F (2005), 'Directed assembly of cylinder-forming block copolymer films and thermochemically induced cylinder to sphere transition: A hierarchical route to linear arrays of nanodots', *Nano Lett*., **5**, 1379–1384.

Liu C-C, Han E, Onses S, Thode C J, Ji S, Goplan P and Nealey P F (2011), 'Fabrication of lithographically defined chemically patterned polymer brushes and mats', *Macromolecules*, **44**, 1876–1885.

Liu G, Thomas C S, Craig G S W and Nealey P F (2010), 'Integration of density multiplication in the formation of device-oriented structures by directed assembly of block copolymer-homopolymer blends', *Adv. Funct. Mater.,* **20**, 1251–1257.

Mansky P, Chaikin P and Thomas E L (1995), 'Monolayer films of diblock copolymer microdomains for nanolithographic applications', *J. Mater. Sci*., **30**, 1987–1992.

Mansky P, Harrison C K, Chaikin, P M, Register R A and Yao N (1996), 'Nanolithographic templates from diblock copolymer thin films', *Appl. Phys. Lett.*, **68**, 2586–2588.

Mansky P, Liu Y, Huang E, Russell T P and Hawker C J (1997a), 'Controlling polymer-surface interactions with random copolymer brushes', *Science*, **275**, 1458–1460.

Mansky P, Russell T P, Hawker C J, Mays J, Cook D C and Satija S K (1997b), 'Interfacial segregation in disordered block copolymers: Effect of tunable surface potential', *Phys. Rev. Lett.*, **79**, 237–240.

Matsui S and Mori K (1984), 'New selective deposition technology by electron beam induced surface reaction', *Jpn. J. Appl. Phys*., **23**, L706–L708.

Matsui S and Mori K (1986), 'New selective deposition technology by electron-beam induced surface reaction', *J. Vac. Sci. Technol. B*, **4**, 299–304.

Matsui S and Ichihashi T (1988), 'In situ observation on electron-beam-induced chemical vapor deposition by transmission electron microscopy', *Appl. Phys. Lett*., **53**, 842–844.

Matsui S, Ichihashi T and Mito M (1989), 'Electron beam induced selective etching and deposition technology', *J. Vac. Sci. Technol. B*, **7**, 1182–1190.

Matsui S, Kaito T, Fujita J, Komuro M, Kanda K and Haruyama Y (2000), 'Three-dimensional nanostructure fabrication by focused-ion-beam chemical vapor deposition', *J. Vac. Sci. Technol. B,* **18**, 3181–3184.

Morkved T L, Lu M, Urbas A M, Ehrichs E E, Jaeger H M, Mansky P and Russell T P (1996), 'Local control of microdomain orientation in diblock copolymer thin films with electric fields', *Science*, **273**, 931–933.

Park M, Harrison C, Chailin P M, Register R A and Adamson D H (1997), 'Block copolymer lithography: Periodic arrays of ~1011 holes in 1 square centimeter', *Science*, **276**, 1401–1404.

Park S-M, Craig G S W, La Y -H, Solak H and Nealey P F (2007a), 'Square arrays of vertical cylinders of PS-b-PMMA on chemically nanopatterned surfaces', *Macromolecules*, **40**, 5084–5094.

Park S-M, Stoykovich M P, Ruiz R, Zhang Y, Black C T and Nealey P F (2007b), 'Directed assembly of lamellae-forming block copolymers by using chemically and topographically patterned substrates', *Adv. Mater.*, **19**, 607–611.

Park S-M, Craig G S W, Liu C-C, Hye Y-H, Ferrier N J and Nealey P F (2008a), 'Characterization of cylinder-forming block copolymers directed to assemble on spotted chemical patterns', *Macromolecules*, **41**, 9118–9123.

Park S-M, Craig G S W, La Y-H, and Nealey P F (2008b), 'Morphological reconstruction and ordering in films of sphere-forming block copolymers on striped chemically patterned surfaces', *Macromolecules*, **41**, 9124–9129.

Rockford L, Liu Y, Mansky P, Russell T P, Yoon M and Mochrie S G J (1999), 'Polymers on nanoperiodic, heterogeneous surfaces', *Phys. Rev. Lett.*, **82**, 2602–2605.

Ruiz R, Kang H M, Detcheverry F A, Dobisz E, Kercher D S, Albrecht T R, de Pablo J J and Nealey P F (2008), 'Density multiplication and improved lithography by directed block copolymer assembly', *Science*, **321**, 936.

Ruiz R, Dobisz E and Albrecht T R (2011), 'Rectangular patterns using block co-polymer directed assembly for high bit aspect ratio patterned media', *ACS Nano*, **5**, 79–84.

Segalman R A, Yokoyama H and Kramer E J (2001), 'Graphoepitaxy of spherical domain block copolymer films', *Adv. Mater.*, **13**, 1152–1155.

Smith H I and Flanders D C (1978), 'Oriented crystal growth on amorphous substrates using artificial surface – relief gratings', *Appl. Phys. Lett.*, **32**, 349–350.

Stoykovich M P, Kang H, Daoulas K C, Liu G, Liu C-C, de Pablo J J, Müller M and Nealey P F (2007), 'Directed self-assembly of block copolymers for nanolithography: Fabrication of isolated features and essential integrated circuit geometries', *ACS Nano,* **1**, 168–175.

Stoykovich M P, Daoulas K C, Müller M, Kang H, de Pablo J J and Nealey P F (2010), 'Remediation of line edge roughness in chemical nanopatterns by the directed assembly of overlying block copolymer films', *Macromolecules*, **43**, 2334–2342.

Suh H, Kang H, Liu C -C, Nealey P F and Char K (2010), 'Orientation of block copolymer resists on interlayer dielectrics with tunable surface energy', *Macromolecules*, **43**, 461–466.

Tada Y, Akasaka S, Yoshida H, Hasegawa H, Dobisz E, Kercher D and Takenaka M (2008), 'Directed self-assembly of diblock copolymer thin films on chemically-patterned substrates for defect-free nano-patterning', *Macromolecules*, **41**, 9267–9276.

Tada Y, Yoshida H, Ishida Y, Hirai T, Bosworth J, Dobisz E, Ruiz R, Takenaka M, Hayakawa T and Hasegawa H (2012), 'Directed self-assembly of POSS containing block copolymer on lithographically defined chemical template with morphology control by solvent vapor', *Macromolecules*, **45**, 292–304.

Takenaka M, Wakada T, Akasaka S, Nishitsuji S, Saijo K, Shimizu H, Kim M I and Hasegawa H (2007), 'Orthorhombic F*ddd* network in diblock copolymer melts', *Macromolecules*, **40**, 4399–4402.

Tokarev I, Krenek R, Burkov Y, Schmeisser D, Sidorenko A, Minko S and Stamm M. (2005), 'Microphase separation in thin films of poly(styrene-block-4-vinylpyridine) copolymer -2-(4'-hydroxybenzeneazo) benzoic acid assembly', *Macromolecules*, **38**, 507–516.

Xiao S, Yang X M, Edwards E W, La Y-H and Nealey P F (2005), 'Graphoepitaxy of cylinder-forming block copolymers for use as templates to pattern magnetic metal dot arrays', *Nanotechnology*, **16**, S324–S329.

Xiao S, Yang X M, Park S, Weller D and Russell T P (2009), 'A novel approach to addressable 4 Teradot/in.^{2} patterned media', *Adv. Mater.*, **21**, 2516–2519.

Zhang X, Berry B C, Yager K G, Kim S and Jones R L (2008), 'Surface morphology diagram for cylinder-forming block copolymer thin films', *ACS Nano*, **2**, 2331–2341.

9

Nanoimprint lithography

D. RESNICK, Molecular Imprints Inc., USA

DOI: 10.1533/9780857098757.315

Abstract: Imprint lithography is essentially a micromolding process in which the topography of a template defines the patterns created on a substrate. When the printing is done at the nanoscale the various imprint methods are generally referred to as nanoimprint lithography (NIL). There are three basic approaches to imprint lithography: soft lithography (SL), thermal nanoimprint lithography (T-NIL), and ultraviolet nanoimprint lithography (UV-NIL). In addition, there is growing interest in a roll to roll (R2R) imprint process for large area substrates such as flat panels. Each approach is described in the following sections. Following this is a discussion devoted to defectivity mechanisms and defect reduction.

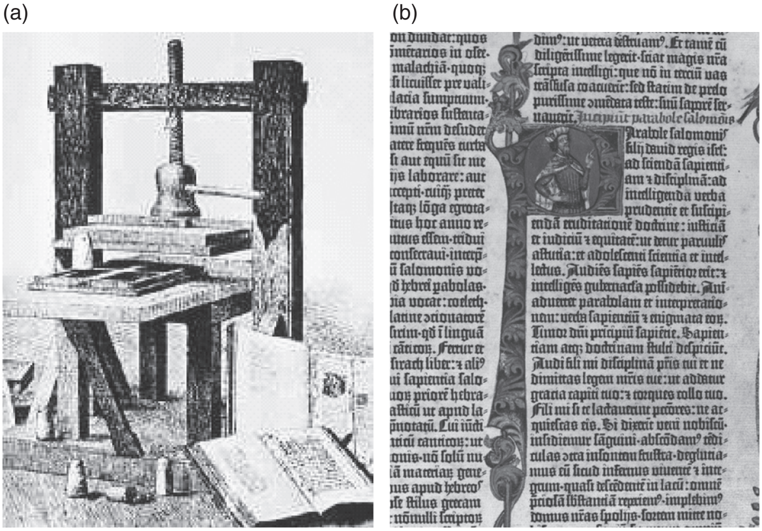
Key words: nanoimprint lithography, NIL, thermal nanoimprint lithography, jet and flash imprint lithography, templates.

9.1 Introduction

Imprint lithography has a long history, and implementation of the technology is now being developed at the nanoscale. Unlike many other lithographic techniques, imprint lithography can be applied cost effectively to address the patterning needs of a diverse number of applications.

9.1.1 The history of imprint lithography

Relative to some of the other lithographic techniques discussed in this book, imprint lithography is a very old and established technology. While Johannes Gutenberg is generally credited with the invention of modern printing, imprinting was already being practiced for centuries in China. As early as 500 BC, there is evidence of carved characters in stone and ceramic. With the invention of paper in approximately 200 BC, it became possible to reproduce works on large writing surfaces. Important work was etched into stone slab and transferred onto materials such as hemp, silk rags, or bark. In the seventh century, stone gave way to woodblock, a technique that was also used in Europe during the Middle Ages. To form a printed image, a sheet of paper was placed over the block, and the image was transferred either by rubbing, or by inking the block and abrading the paper.



9.1 (a) The Gutenberg press. (b) A page from the 42 Line Bible.

Wood block remained the primary means of imprinting well into the nineteenth century in China. However, in 1040, Bi Sheng began to experiment with movable type made from either clay or ceramics. The individual type was arranged on an iron plate, and held in place with wax and resin. By melting the wax, the type could be removed and then rearranged as necessary. Eventually, clay gave way to wood, and later to more durable materials such as copper and brass. The technology never became popular in China, because of the thousands of characters that were needed to convey information in the Chinese language.

Johannes Gutenberg began his work in the town of Mainz in 1436. His invention was a combination of the formation of easily cast characters combined with a screw type press, which was a takeoff of machines that were used to produce wine in the Rhine valley. By splitting individual components of language such as letters, numbers, and punctuation into individual units, it was now possible to quickly form different sentences and paragraphs. The press was completed in 1440, and in 1455 printed versions of a 42 Line Bible (42 lines per page) appeared. Two hundred copies of this Bible were printed, and forty-eight copies are known to exist today. A picture of the press, along with a page from the Bible, is shown in Fig. 9.1.

Not unlike start-ups today, Gutenberg's work required venture capital, and a moneylender by the name of Johannes Fust realized the potential of Gutenberg's work. In 1449, Fust loaned Gutenberg 800 florins, which was used for the preparation of the imprinting tool (http://www.uh.edu/engines/epi753.htm). Prior to the release of the Bible, several books and treatises were published, and the impact of the technology began to blossom. By the time of the printing of the Bible, however, the relationship between

Gutenberg and Fust had become strained, and Fust took Gutenberg to court, claiming, among other things, embezzlement of funds. The archbishop's court of worldly justice ruled primarily in favor of Fust, and in the same year that the 42 Line Bible was printed, Gutenberg lost his printing workshop, and was effectively bankrupt.

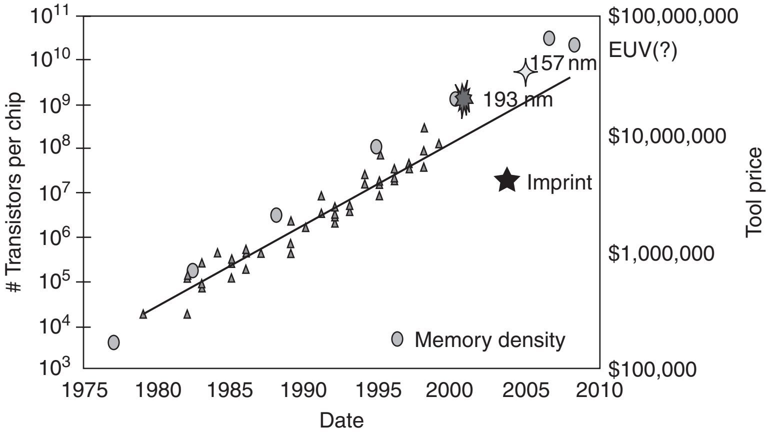
Gutenberg continued his work, along with other printers, up until 1462, when the Archbishop of Nassau attacked the city of Mainz. Once the city was under control, the surviving citizens, including Gutenberg, were forced to leave. When printers and compositors settled in new towns, the knowledge of the printing process was spread beyond the immediate vicinity of Mainz. By 1477, William Caxton published the first book in England. By the end of the century, printing technology was established in over 250 cities across Europe. Roman type was introduced in 1572, and Oxford University started a printing operation in 1587. In 1593 Shakespeare's *Venus and Adonis* appeared in print and began a new era in literature.

It should also be noted that although movable type is generally thought of as a European invention, a common use of movable type began in Korea around the same time as Gutenberg's innovative work. An alphabetical script known as 'Han'gul,' originally consisting of 28 characters, was officially presented in 1444.

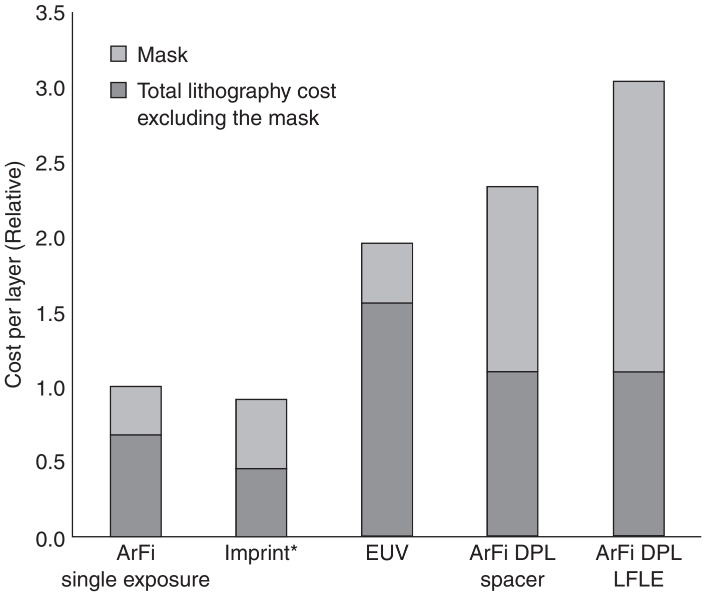
9.1.2 The rising cost of lithography

High-density semiconductor circuits are now being mass produced with critical dimensions (CD) less than 30 nm. In the fields of micro- and nano-lithography, major advancements in resolution have historically been achieved through use of shorter wavelengths of light. Using phase shift mask technology, it has already been demonstrated that 193 nm photolithography can produce sub-100 nm features. By applying double patterning processes to immersion-based 193 nm systems, devices such as NAND Flash memory are in production at half pitches as small as 19 nm. Along this path, such improvements come with an ever-increasing cost for photolithographic tools. It is interesting to note that the cost of photo-tools has kept pace with data density, as depicted in Fig. 9.2.[1] As conventional projection lithography reaches its limits, next generation lithography (NGL) tools, such as extreme ultraviolet lithography (EUVL), may provide a means for further pattern shrinks, but are expected to have a price tag that is prohibitive for many companies.

The development of both light sources and optics to support the sources is primarily responsible for the rise in the cost of an NGL tool. 157 nm lithography (a technology that never made it to production), for example, required the use of CaF_2 as a lens material. In the case of EUVL, no source with sufficient output has yet been identified that will meet the industry's throughput requirements.



9.2 Transistor density and tool cost as a function of time.

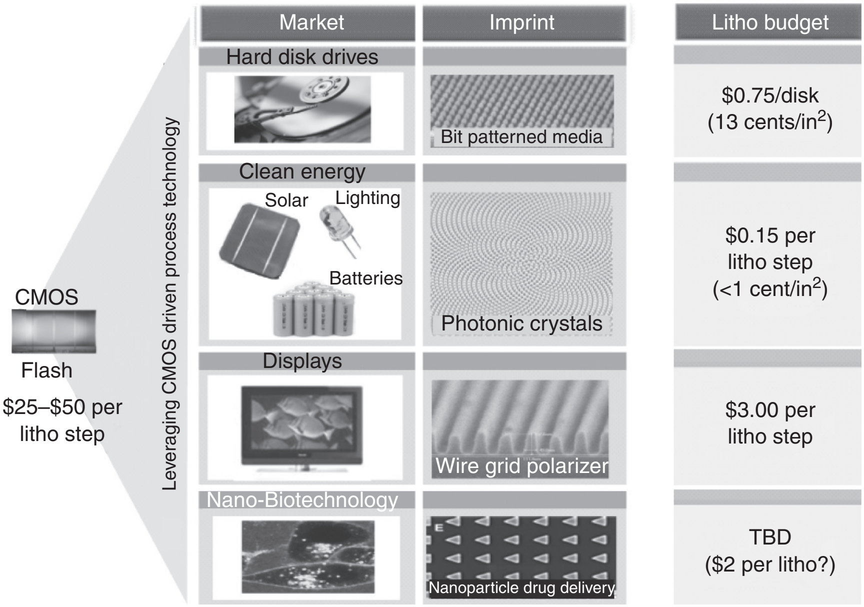


9.3 Cost of ownership at 22 nm for different lithographic processes. The tool and mask costs play a key role in determining cost of ownership. Source data: SEMATECH; SPIE advanced lithography, 2009.
*Mil mask data including replication

Several models exist that can estimate the cost of ownership (or cost per wafer level). Two key ingredients in the model are tool cost and mask cost. Figure 9.3 is a plot of cost of ownership for competing lithographic processes for 22 nm at a throughput of 20 000 wafers per month.[2] Clearly, an imprint technology that can reduce the tool cost by an order of magnitude will have a significant effect on the economics of the fabrication process.

The unique physical and chemical phenomena at the nanoscale can lead to other types of novel devices that have significant practical value. Emerging nano-resolution applications include sub-wavelength optical components, biochemical analysis devices, high-speed compound semiconductor chips, distributed feedback lasers, photonic crystals, and high-density patterned magnetic media for storage. To take advantage of these opportunities, it is necessary to be able to cost effectively image features well below 100 nm.

It is interesting to compare the cost of lithography for a high-density semiconductor device, such as NAND Flash memory, with some of the applications mentioned in the previous paragraph. As illustrated in Fig. 9.4, the lithography cost per wafer level for a complementary metal- oxide semiconductor (CMOS) NAND Flash critical layer is somewhere between \$25 and \$50. Other markets, including patterned media for hard drives, clean energy, displays, and bio-medicine have comparable patterning needs, but require reductions in lithography costs by as much as three orders of magnitude.[3] This can be accomplished only with high resolution patterning tools that deliver both


Market
Imprint
Litho budget
Hard disk drives
Bit patterned media
$0.75/disk
(13 cents/in²)
Clean energy
Solar
Lighting
Batteries
Photonic crystals
$0.15 per
litho step
(<1 cent/in²)
Displays
Wire grid polarizer
$3.00 per
litho step
Nano-Biotechnology
Nanoparticle drug delivery
TBD
($2 per litho?)
CMOS
Flash
$25–$50 per
litho step
Leveraging CMOS driven process technology

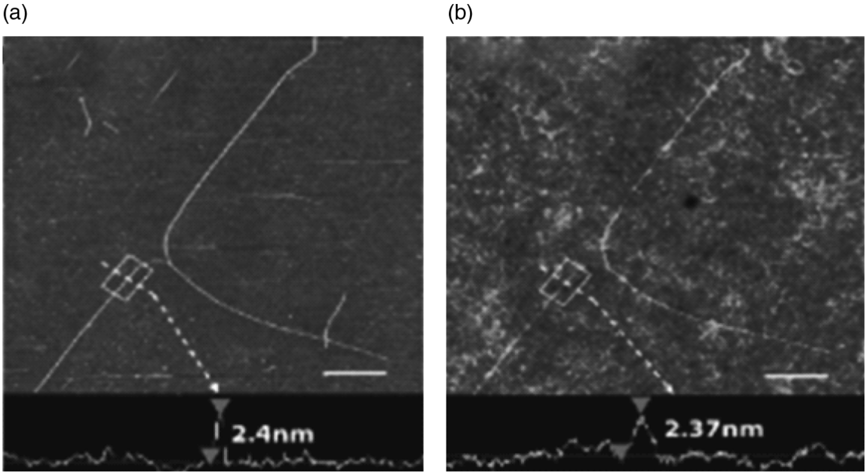
9.4 Cost of ownership for several markets requiring high resolution patterning. Relative to semiconductor, the other markets have much smaller lithography budgets.

lithographic performance and throughput, as well as minimizing the material costs of the process. This is one of the unique benefits of imprint lithography.

9.2 An overview of imprint lithography

Imprint lithography is essentially a micromolding process in which the topography of a template defines the patterns created on a substrate. When the printing is done at the nanoscale the various imprint methods are generally referred to as NIL. Investigations by several researchers indicate that imprint lithography resolution is limited only by the resolution of the template fabrication process. The most famous example of this capability is the reproduction of a 2 nm carbon nanotube by the John Rogers Research group at the University of Illinois (Fig. 9.5).[4] NIL possesses other important advantages over conventional photolithography and other NGL techniques, since it does not require expensive projection optics, advanced illumination sources, or specialized resist materials that are central to the operation of these technologies.

There are three basic approaches to imprint lithography: soft lithography (SL), thermal nanoimprint lithography (T-NIL), and UV-NIL. In addition, there is growing interest in an R2R imprint process for large area substrates, such as flat panels. Although the imprint process is not unique in itself, the techniques used to create nanopatterns from a roller are different enough to warrant a separate discussion. Each approach is described in the following sections. Following this is a discussion devoted to defectivity mechanisms and defect reduction.

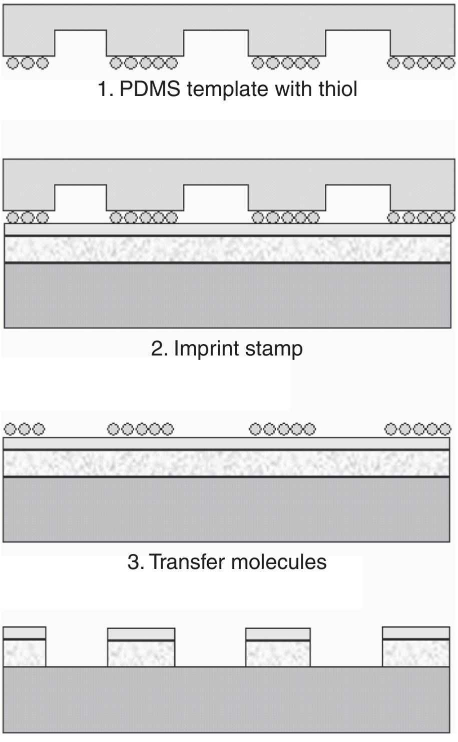


9.5 (a) A master mask consisting of deposited 2.4 nm diameter carbon nanotube. (b) An imprint of the carbon nanotube.

9.3 Soft lithography

SL, also known as microcontact printing (μCP), generally refers to the process of transferring a self-assembled monolayer using a flexible template. The invention of the technology dates back to 1994, and is the result of work from the laboratory of George Whitesides at Harvard.[5,6] The technology surfaced mainly as a quick and easy way for students to print small geometries in a laboratory environment. Whitesides *et al.* formed a template by applying a liquid precursor to polydimethylsiloxane (PDMS) over a master mask produced using either electron beam or optical lithography. The liquid is cured, and the PDMS solid is peeled away from the original mask. The PDMS is essentially an elastomeric material, consisting of a polymer chain of silicon containing oils. Typical mechanical properties include a tensile strength of 7.1 MPa, an elongation at break of 140%, and a tear strength of 2.6 kN/m. As a result, relative to either silicon or fused silica, it is quite pliant.

Once prepared, the PDMS template can then be coated with a thiol ink solution, such as an alkanethiol. The imprint process is depicted in Fig. 9.6.



9.6 Process schematic for SL.

The thiol molecules are subsequently transferred to a substrate, coated with a thin layer of gold, thereby forming a self-assembled monolayer (SAM) on the gold surface. The nature of the gold-sulfur bond is still not completely understood. R. S. Kane *et al.* postulate that the species present at the surface of the gold is a gold thiolate[6]:

$$\text{R-SH} + \text{Au}(0)_n \rightarrow RS - \text{Au (I) Au } (0)_{n\text{-}1} + \rightarrow\uparrow \frac{1}{2}\text{H}_2 \rightarrow\uparrow \qquad [9.1]$$

To prevent adhesion between the master and daughter masks, the master surface is passivated by the gas phase deposition of a long-chain, fluorinated alkylchlorosilane ($CF_3(CF_2)_6(CH_2)_2SiCl_3$). The fluorosilane reacts with the free silanol groups on the surface of the master to form a Teflon-like surface with a low interfacial free energy. The passivated surface acts as a release layer that facilitates the removal of the PDMS stamp from the master.

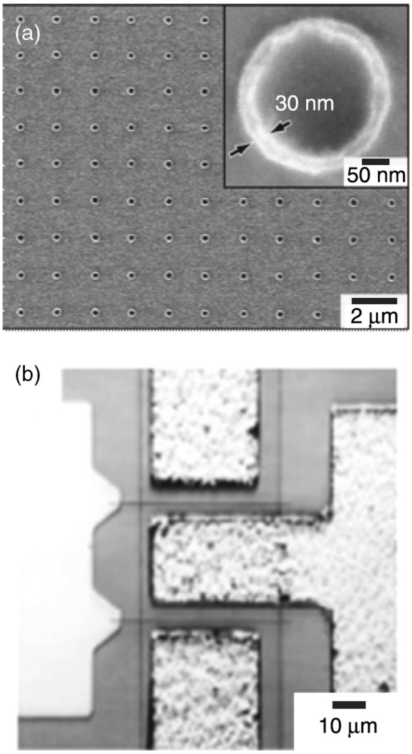
The pattern transfer process starts with a wet etch of the thin gold film. A wet etch is typically used, since gold is not readily reactively ion etched. Although it is possible to sputter etch gold, the thin thiol layer would not hold up to such a process. The gold film then acts as an etch mask for any underlying materials. Because gold is a soft metal, it is often necessary to include a second hard mask beneath the gold.

The range of feature sizes that can be imprinted with this technology is quite broad. While squares and lines with geometries of several microns are easily achieved, smaller circular features with dimensions as small as 30 nm have also been demonstrated (Fig. 9.7).

Although the technology has been used to make working field effect transistors, complex electronic devices are not likely to be the strength of this particular printing technology. Although the SAMs are easily transferred and tend to self-heal during deposition, the thickness of the molecule (~1 nm) makes it difficult for routine pattern transfer of the thicker films typically required for semiconductor devices. Thinner films are also subject to defects, and yield is critical in the semiconductor field. In addition, the elastic nature of the PDMS template makes it impractical for the very precise layer-to-layer overlay tolerances required in the semiconductor industry. Interestingly enough, however, the same feature that renders it difficult to align one level to another also allows printing on surfaces that are not planar. This attribute is unique to microcontact printing.

9.4 Thermal imprint lithography

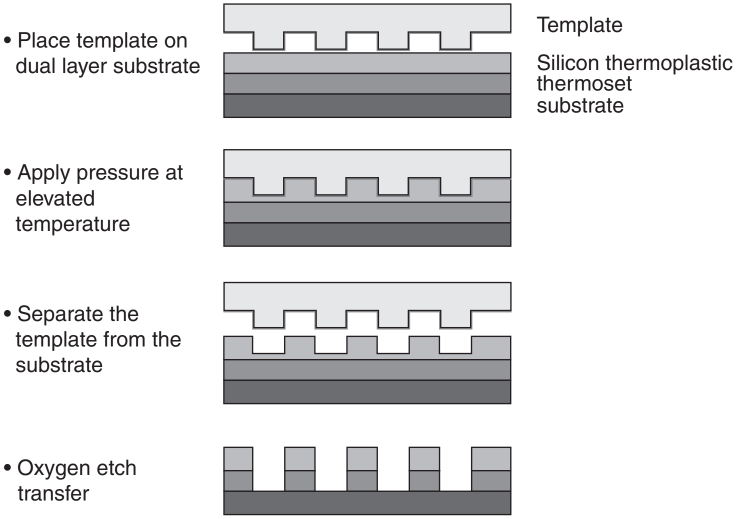
Thermal imprint lithography was first introduced by Dr. Stephen Chou from the University of Minnesota in 1996.[7] Chou introduced a mechanical



9.7 (a) A 30 nm ring created with µCP. (b) A working field effect transistor.

printing process that did not require any type of energetic beam. A schematic of the imprint process is depicted in Fig. 9.8. A resist, such as poly(methyl methacrylate) (PMMA), is spin coated and baked onto a substrate, such as silicon. The substrate and resist are then heated above the glass transition temperature of the resist. For PMMA, a typical imprint bake temperature is between 140°C and 180°C. The template is then pressed into the resist until the resist flows into the features of the mold. Typical imprinting pressures for the initial experiments range from 600 to 1900 psi (40–129 atmospheres!). Pressures of less than 30 atmospheres are more typical of today's processes. The template and substrate are then cooled, and separated, leaving a reverse tone image that has the potential to replicate precisely the features on the mold.

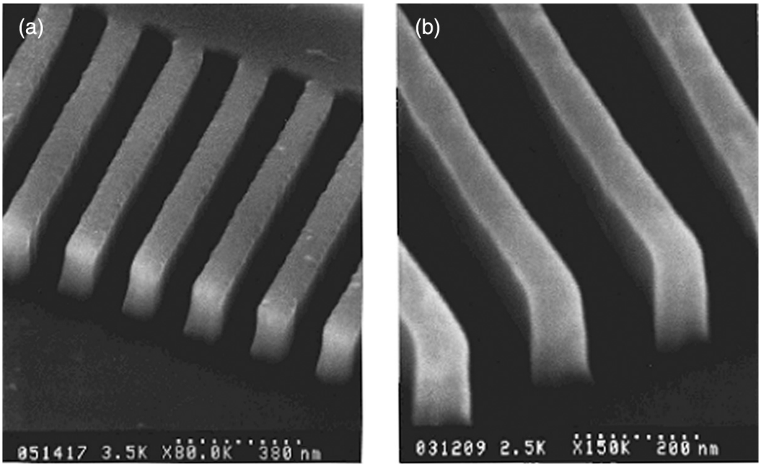
It is possible to use a variety of different resist materials, and several materials are now manufactured specifically for use with imprint lithography. PMMA was initially chosen for its low thermal coefficient of expansion and small shrinkage coefficient. Thermal imprint lithography is usually carried out in a vacuum environment in order to avoid trapping air in the


• Place template on dual layer substrate
Template
Silicon thermoplastic thermoset substrate
• Apply pressure at elevated temperature
• Separate the template from the substrate
• Oxygen etch transfer

9.8 Schematic flow for a thermal imprint process.

template during the imprint process. 'Sticking' issues are mitigated either by adding release agents to the resist or by treating the surface of the template with a low surface energy material such as a 1H, 1H, 2H, 2H-Perfluorooctytrichlorosilane (FOTS) anti-adhesive monolayer.[8]

The template used in the imprinting process is also a silicon substrate. This is a necessity, since the NIL takes place at elevated temperatures, and care must be taken to avoid thermal mismatches during imprinting. There are two common ways to pattern the silicon. The first simply requires the use of a resist mask to etch the underlying silicon. For sub-100 nm feature definition, it is necessary to use a high resolution electron beam writing tool. These tools typically have one of two form factors: (1) For best resolution, Gaussian, or spot beam tools, can be used to form resist patterns as small as 10 nm. Because all of the exposure area uses the same small spot size, write times for complex patterns can be as long as several days. (2) Fastest writing times are achieved with shaped beam tools, where the ultimate resolution is dictated by the size of the aperture and the electron beam blur from the aperture. Dimensions as small as 22 nm have been demonstrated, with feature image placements as low as 3 nm. Alternatively, scanning electron microscopes, equipped with customized e-beam writing packages, can provide excellent resolution. Typical high resolution electron beam resists include PMMA and ZEP520A. Where it is necessary to etch deep into the silicon, a hard mask may be employed in the pattern transfer process. Alternatively, a metal lift-off process may follow e-beam imaging. The remaining metal then serves as an etch mask for the underlying silicon. The silicon can be dry



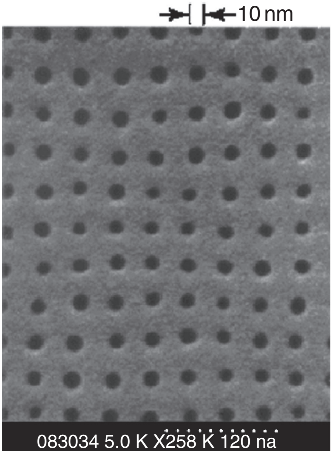
9.9 70 nm features in the mold (a), and in the imprinted resist (b).

etched with either chlorine- or bromine-containing gases. These chemistries can be quite aggressive on electron beam resists. As a result, some type of hardmask, as described above, may be required.

A second method for forming a silicon template involves the deposition of a thin oxide film on the silicon, followed by electron beam imaging. The e-beam resist then serves as a mask for the oxide, and the silicon acts as an etch stop for the oxide. This is a very common method now used for defining thermal imprint templates. Very little data have been published on the specifics of this template fabrication process, but the oxide etch process itself is well known. Details on the etching characteristics of fused silica are discussed later in this chapter.

Agreement between the template image and the final resist image can be quite remarkable. A comparison between 70 nm template and resist features is depicted in Fig. 9.9. In Chou's first work on the subject, 60 nm lines were imaged in PMMA resist. This is clearly not the limit of resolution of the technology, however. Subsequent work has demonstrated that it is possible to print holes smaller than 10 nm (see Fig. 9.10), again demonstrating that resolution of the technology is limited only by what can be achieved in the template.[9]

The imprinted image is not useful until the pattern transfer process is completed. Any molding type of imprint process is typically accompanied by a residual layer after the imprint process, as shown in Fig. 9.7. The first step in the pattern transfer process, therefore, involves a descum process to remove the residual layer. For PMMA, an oxygen descum works quite well; however, care must be taken to control this step, in order not to introduce a lateral etch component, thereby changing the CD of the feature of interest. Once this step is complete, either a subtractive or additive process can

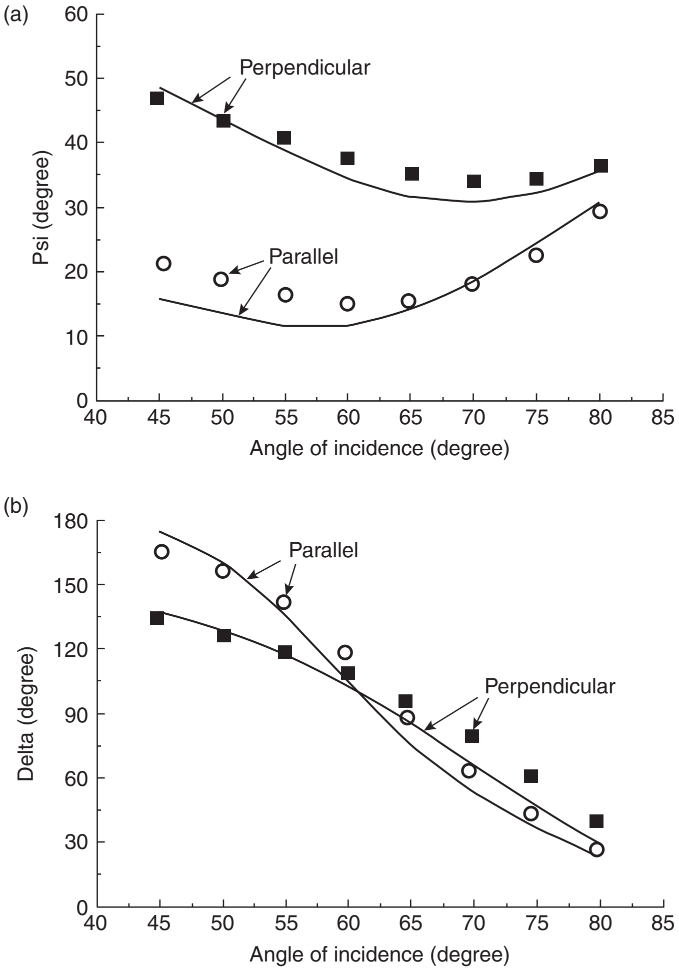


9.10 Sub-10 nm holes formed using T-NIL.

occur. PMMA has notoriously poor etch selectivity. As a result, most of the early work tended to focus on lift-off processing.

Because the thermal imprint process requires no optics, it is easily implemented in research labs and has been used to fabricate a variety of interesting and useful devices. Two examples are discussed below. Interested readers are referred to References 10–14 for further examples of fabricated devices. Gratings are relatively easy to print using thermal imprint lithography, and may be used for a variety of interesting applications, such as extreme-ultraviolet and ultraviolet filters, transmission filters, visible and infrared polarizing optical elements, waveplates, and phase retarders. Yu *et al.* formed 100 nm pitch grating in a silicon mold by first patterning oxide features on a 200 nm pitch.[14] By depositing, silicon nitride, performing an etch back, removing the oxide, and using the nitride as an etch mask into the silicon, he was able to form large area 100 nm period gratings. The usefulness of the resultant gratings was tested by measuring the ellipsometric parameters of the beam at a wavelength of 632.8 nm at different angles of incidence, with the gratings oriented either parallel or perpendicular to the plane of incidence. Good agreement is obtained with effective medium theory simulations, as shown in Fig. 9.11.

Microring resonator devices have been the focus of recent interest because of their potential for applications in photonic circuits.[15] A microring resonator typically has the shape of a ring closely coupled to a waveguide, thereby offering capabilities such as narrow bandwidth filtering and compactness. Chao *et al.* studied two methods for fabricating polymer-based microring resonators.[16] In one method, T-NIL was used to imprint a PMMA waveguide on top of thermal oxide. The PMMA was heated to a temperature of 175°C,

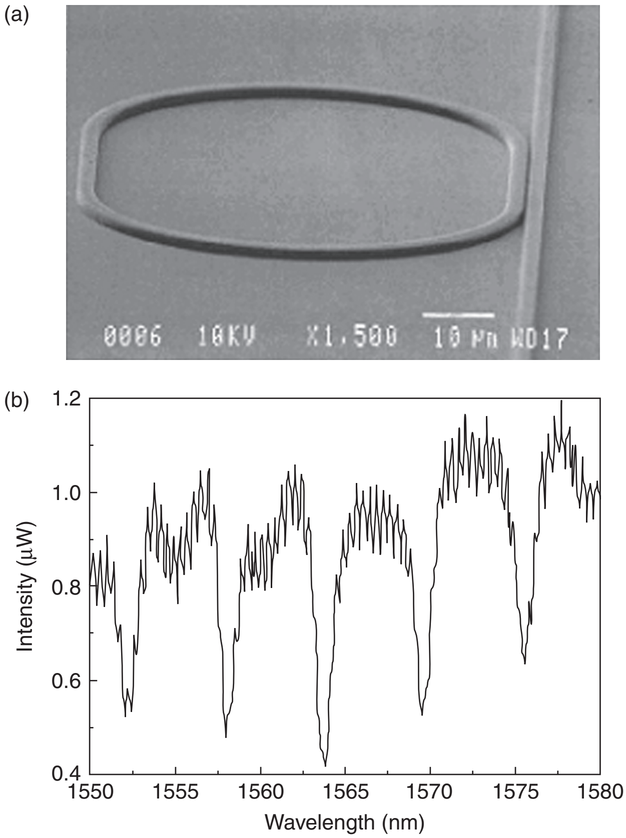


9.11 (a) Parametric variance of the beam at a wavelength of 632.8 nm at different angles of incidence, with the gratings oriented either parallel or perpendicular to the plane of incidence. (b) Good agreement is obtained with effective medium theory simulations, shown as a solid line.

and the imprint was performed at a pressure of 75 kg/cm^2. Figure 9.12a shows the oxide mold used for the imprint process. Figure 9.12b shows the output of a fabricated PS microring resonator

9.5 Alternative thermal imprint processes

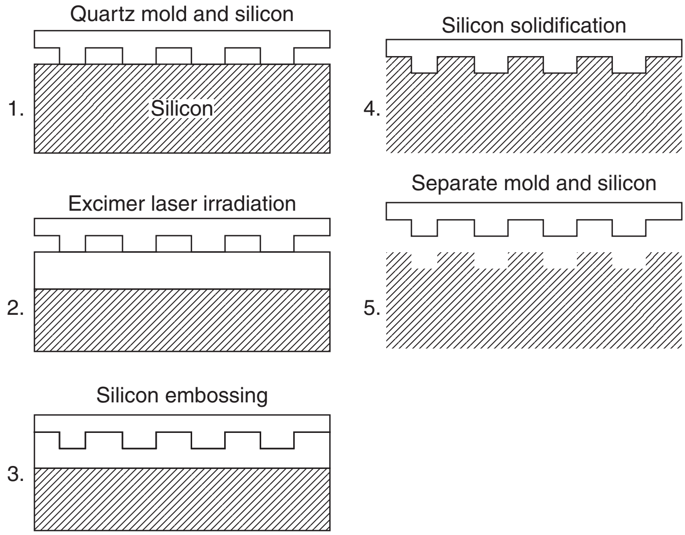
All of the lithographic methods discussed upto this point, including the more conventional forms of photolithography, involve the use of an imaging



9.12 (a) Microring resonator mold. (b) Output spectra response with a TM-polarized input.

resist, which is then utilized as a mask to define an image into an underlying material. Any type of pattern transfer process is problematic, since it has the potential for adding defects and changing CDs. It would be very advantageous to be able to directly image the material of choice without the need of a resist. The development of functional materials is under way, and dielectric materials, such as hydrogen silsesquioxane, have been proved to be both electron beam- and photo-sensitive.[17] Imprint lithography combined with functional materials would be a very cost-effective way of patterning features.

Chou *et al.* have developed a variant of thermal imprinting to pattern silicon directly. The technique is called laser-assisted direct imprint (LADI).[18] A schematic of the LADI process is shown in Fig. 9.13. A quartz mold is used to form the features in the silicon. To image the silicon, a 20 ns XeCl

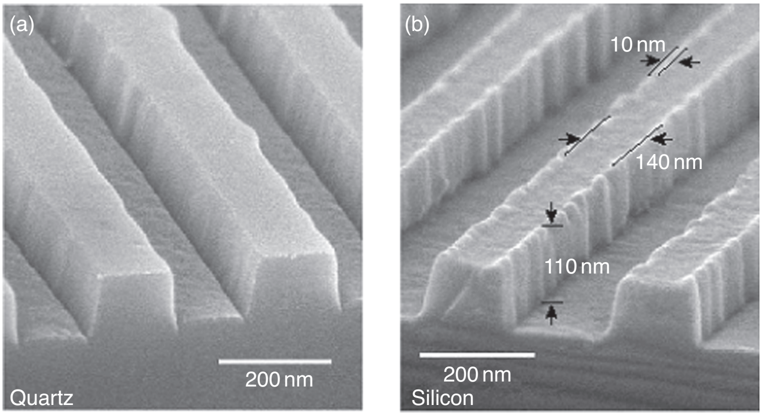


9.13 Schematic representation of the LADI method for patterning silicon.

laser pulse (at a wavelength of 308 nm) is applied through the quartz template and onto the silicon surface. The energy imparted by the laser causes a thin layer on the surface of the silicon to melt. Because the viscosity of the silicon melt is extremely low, the quartz can then be pressed into the silicon, allowing the silicon melt to fill the mold. It was estimated that the silicon remains as a melt for approximately 220 ns and that the depth of the melt is roughly 280 nm. Once the silicon cools, the quartz and silicon can be separated.

As a demonstration, Chou patterned 300 nm grating structures, 110 nm deep into silicon. As with conventional NIL, the patterns were faithfully replicated. Scanning Electron Microscope (SEM) micrographs of both the mold and the replicated features are depicted in Fig. 9.14. Although the exact force of the imprint process could not be measured, it is estimated that the force used for the small demonstration pieces was approximately 17 atmospheres. It is possible that the process, which was performed on 1.5 mm × 1.5 mm samples, could be scaled up to sizes comparable to those used in conventional NIL tools.

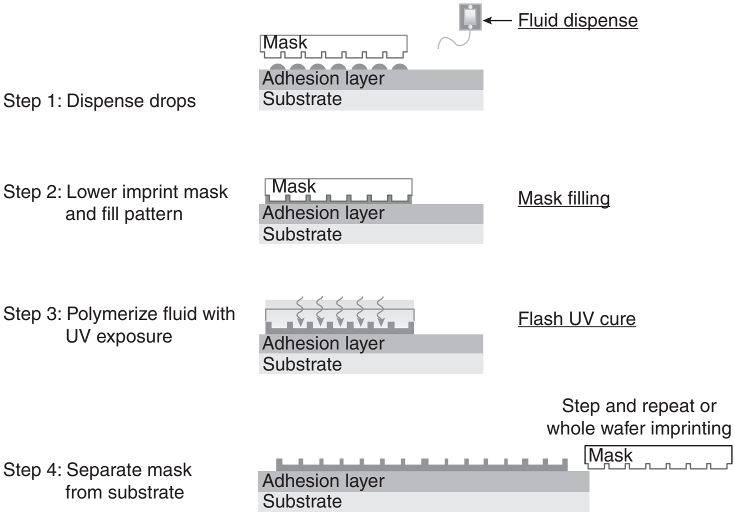
It should be noted that LADI is not limited only to patterning silicon. Polysilicon patterning has also been demonstrated, but it may be feasible to extend the technology to very different materials such as germanium, III-V compounds, and dielectrics. Three dimensional patterning should also be possible with the technique. Future development of this exciting technology will be interesting to monitor.



9.14 SEM image of the template (a), and of the patterned silicon (b), using LADI.

9.6 Ultraviolet (UV) nanoimprint lithography overview

Devices that require several lithography steps and precise overlay will need an imprinting process capable of addressing registration issues. A derivative of NIL, ultraviolet nanoimprint lithography (or UV-NIL), addresses the issue of alignment by using a transparent template, thereby facilitating conventional overlay techniques. In addition, the imprint process is performed at low pressures and at room temperature, which minimizes magnification and distortion errors. Two types of approaches are being considered for UV-NIL. The first method uses conventional spin-on techniques to coat a wafer with a UV curable resist.[19] Although it is possible to coat the wafer uniformly, there are concerns that the viscosity of the resist will be too high to facilitate the formation of very thin residual layers. If the residual layer is too thick, the CD uniformity may suffer as a result of the subsequent pattern transfer process. In addition, a uniform coating of resist cannot account for variations in pattern densities on the template or mask, thereby leading to non-uniform residual layers. This problem is addressed by locally dispensing a low viscosity resist. This second approach was first disclosed by Willson *et al.* in 1999 and is generally referred to as step-and-flash imprint lithography (S-FIL) (Fig. 9.15).[20] Advances in S-FIL have led to inkjet based dispensers with drop volumes as small as 1.0 picoliters. This inkjetting approach has been successfully employed for both semiconductor applications where a step and repeat approach is required, and to large area imprinting where throughput demands a process using a single-imprint step to pattern the entire substrate. Because both applications use inkjetting


Fluid dispense
Mask
Adhesion layer
Substrate
Step 1: Dispense drops
Step 2: Lower imprint mask and fill pattern
Mask
Adhesion layer
Substrate
Mask filling
Step 3: Polymerize fluid with UV exposure
Adhesion layer
Substrate
Flash UV cure
Step and repeat or whole wafer imprinting
Mask
Step 4: Separate mask from substrate
Adhesion layer
Substrate

9.15 Schematic illustration of the step and flash imprint lithography (S-FIL) process.

to deposit the resist, the technology has been rebranded as jet and flash imprint lithography (J-FIL).

9.7 Jet and flash imprint lithography

The three primary building blocks that contribute to a UV imprint lithography process are the imprint tool, imprint materials, and an imprint mask. Because the tool and material are so intimately connected during the imprint process, they are discussed as one subject in the next section. Following this discussion, imprint mask attributes and performance results are described.

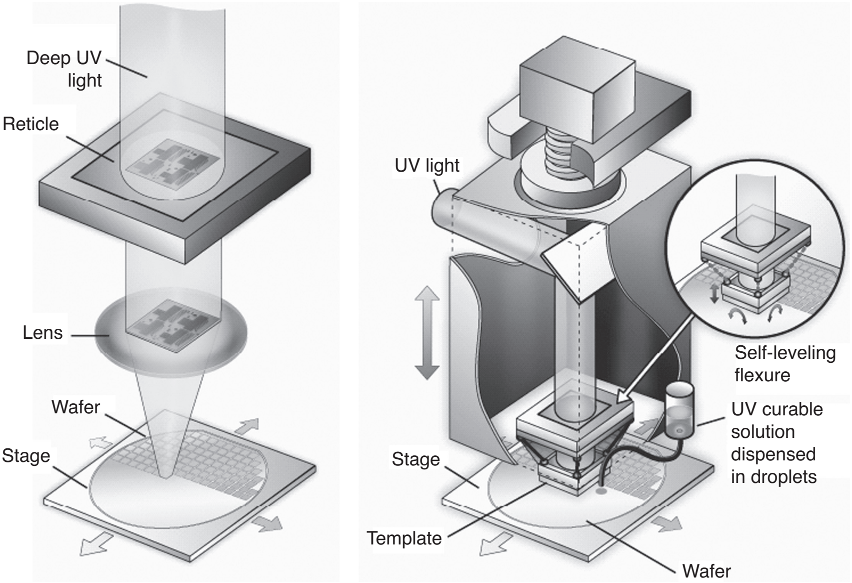
9.7.1 The imprint tool

Imprint tools based on drop-on-demand UV nanoimprinting are broadly divided into wafer steppers for applications requiring nano-resolution overlay and mix-and-match with photolithography (such as silicon intergrated circuit (IC)s and thin film heads for magnetic storage) wherein the field size printing in one patterning step is the same as the industry standard advanced photolithography field size (26 mm × 33 mm), and whole substrate tools for applications that do not require nano-resolution overlay (such as patterned media and photonic crystals for LEDs). Because the material is dispensed only where needed just prior to patterning, the basic drop-on-demand process shown in Fig. 9.15 can be integrated into either tooling platform.

The J-FIL process relies on photo-polymerization of a low viscosity, acrylate-based monomer solution. Acrylate polymerization is known to be accompanied by volumetric shrinkage, which is the result of chemical bond formation. Consequently, the size and shape of the replicated features may be affected. Volumetric shrinkage is found to be less than 10% (v/v) in most cases.[21] Most acrylate-based imprint fluids that have been used to-date have a viscosity in the range of 4–12 cps.[22]

Relative to the spin-on UV-NIL approach, the J-FIL stepper has unique cost advantages. In spin-on UV or thermal imprint processes, a separate spin coating tool is needed for resist deposition. Capital cost in a J-FIL tool is controlled by the inclusion of a self-contained material dispense module. With respect to the cost of consumables, the drop–dispense approach has virtually no waste. It is estimated that the drop–dispense approach consumes 1–0.1% of the volume that a spin coating process consumes. Because semiconductor fabrication requires highly purified materials with ppb contamination levels, the cost of these materials is high and is generally proportional to the volume of the material used.

Figure 9.16 shows the schematic of a photolithography stepper and a UV imprint stepper based on drop–dispense of the imprint material. The imprint tools include precision self-leveling flexure systems to passively align the imprint mask and substrate to be substantially parallel during the imprint process.[23] In addition, by using a drop–dispense approach that can



9.16 Comparison between an optical projection stepper and a J-FIL stepper.

be tailored based on mask pattern variation, a highly uniform residual layer can be achieved. This film needs to be thin and uniform to achieve a subsequent etch process with a high degree of long-range CD uniformity.

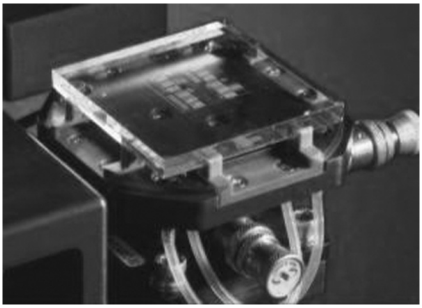
The stepper system has additional precision mechanical systems to achieve nano-resolution alignment and overlay. The alignment sub-system that aligns the imprint mask in x, y, and theta directions with respect to the wafer is based on a field-to-field moiré detection alignment scheme, originally developed at Massachusetts Institute of Technology (MIT) for X-ray lithography.[24] In addition to alignment, magnification and shape corrections are required to perform nano-resolution overlay, particularly when mixing-and-matching to optical lithography projection tools. A precision mechanical deformation system that deforms the imprint mask has been developed and implemented as part of the stepper system.[25,26] This is the key step that allows the stepper to achieve the overlay required for CMOS fabrication.

9.7.2 The imprint mask

Early semiconductor mask fabrication schemes started with a 6″ × 6″ × 0.25″ (6025) conventional photomask plate and used established Cr and phase shift etch processes to define features in the glass substrate.[27] Although sub-100 nm geometries were demonstrated, CD losses during the etching of the thick Cr layer etch make the fabrication scheme impractical for 1X templates. It was not unusual, for example, to see etch biases as high as 100 nm.[28]

More recently, a much thinner (<15 nm) layer of Cr has been used as a hard mask. Thinner layers still suppress charging during the e-beam exposure of the template, and have the advantage that CD losses encountered during the pattern transfer through the Cr are minimized. Because the etch selectivity of glass to Cr is better than 18:1 in a fluorine-based process, the Cr layer is also sufficient as a hard mask during the etching of the glass substrate. Other mask fabrication schemes, including the incorporation of a conductive and transparent layer of indium tin oxide (ITO) on the glass substrate, have also been tested, but the process flow is not easily compatible with the infrastructure available in commercial mask shops. The experimental details of this alternative approach have been covered in previous publications.[29,30]

An example of a patterned J-FIL imprint mask is shown in Fig. 9.17. A mesa defines the field size and is set to have maximum dimensions of 26 mm × 33 mm, thereby making it compatible with existing optical projection lithography tools. Typical mesa heights are on the order of 15–30 μm. The mesa height is critical, as it also provides an impediment to any liquid resist extruding at the edge of the field.



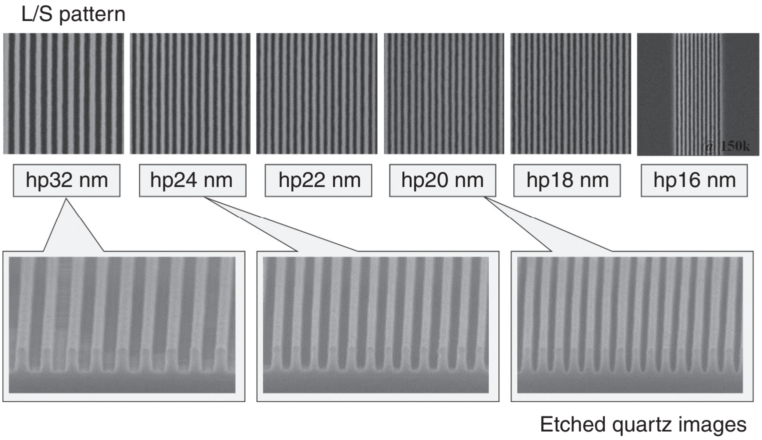
9.17 Photograph of a J-FIL mask.

As previously discussed, imprint resolution is defined by the relief image in the mask. Early work at Motorola took advantage of the high resolution of a Gaussian beam writer and ZEP520A resist to produce imprinted features as small as 30 nm.[31] Resist patterning is followed by a descum step, a dry Cr etch (using chlorine and oxygen), and a reactive ion etch of the fused silica (with chemistries such as CHF_3 or CF_4). More recently, Dai Nippon Printing has demonstrated 14 nm half pitch patterns on a mask. Although impressive, Gaussian beam writers suffer from long write times. Commercial mask suppliers typically prefer to limit write times to less than 24 h, and therefore use shaped beam systems in which an aperture is employed to define minimum resolution. Dense 22 nm line patterns have been resolved using a NuFlare EBM7000.[32] Further refinements in the electron optics, combined with a reduction in the Coulomb interaction, should result in sub-20 nm feature resolution. An example of Dai Nippon Printing (DNP)'s mask fabrication capability is shown in Fig. 9.18.

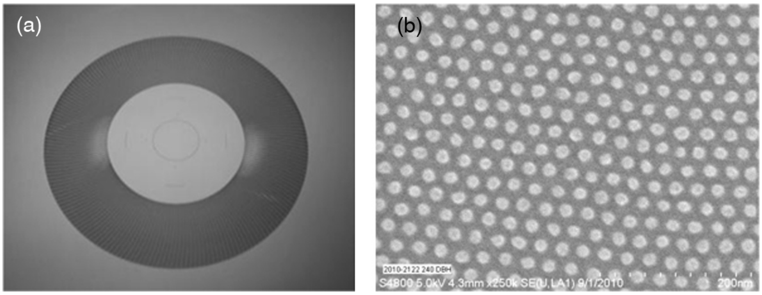
In addition to resist patterning and dry etch pattern transfer, the final mask must go through inspection and repair steps in order to identify and eliminate defects. These steps are typically accomplished using electron beam inspection and repair equipment. References 33, 34, and 35 cover the work that has been done in this field.

For patterning of hard drive media a 6 in. diameter fused silica wafer is used as the mask (template) blank. The blank is patterned in the same way as a semiconductor mask; however, rotary stage Gaussian beam writers are used for the exposure of the resist.[36] Bit patterns, typically referred to as bit patterned media (BPM), are required to achieve aerial densities greater than 1 Tb/in^2, and have respective pitches of less than 25 nm. An example of a patterned media template and an image of a 1 Tb patterned disk is shown in Fig. 9.19.

Mask lifetime is dictated mainly by the interaction of particles with the mask. It is anticipated that lifetimes will fall somewhere between 10 000 and 100 000 imprints, thereby creating a need for replica masks (or templates).


L/S pattern
@ 150k
hp32 nm
hp24 nm
hp22 nm
hp20 nm
hp18 nm
hp16 nm
Etched quartz images

9.18 Top down and cross section SEM images of an imprint mask for half pitches ranging from 16 to 32 nm.


(a)
(b)
2010-2122 240 DBH
S4800 5.0kV 4.3mm x250k SE(U,LA1) 9/1/2010
200nm

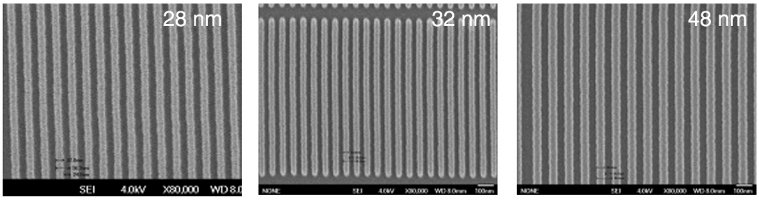
9.19 (a) Photograph of a full field patterned media template. (b) Imprinted bit patterns with a pitch of 25 nm.

Replicas can be fabricated using the same type of imprinting tools used for wafer or disk imprinting.[37,38] An example of a replicated mask pattern is shown in Fig. 9.20.

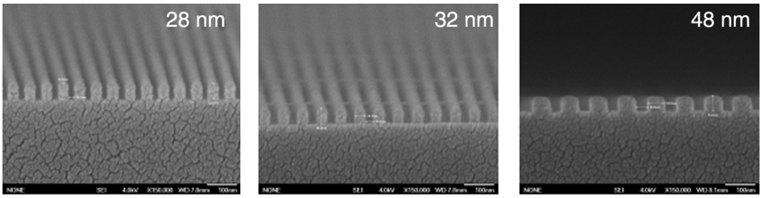
9.7.3 J-FIL performance results

Other than defectivity, the primary attributes that must be addressed are CD control, residual layer control, overlay, and throughput.

While CD uniformity control is required for any lithographic system, the CD is typically a function of mask resolution, and there is very little contribution to the CD error budget as it relates to the actual printed feature. Examples of high resolution patterning are shown in Fig. 9.21. Residual layer thickness (RLT) control is critical, however, for a couple of reasons;

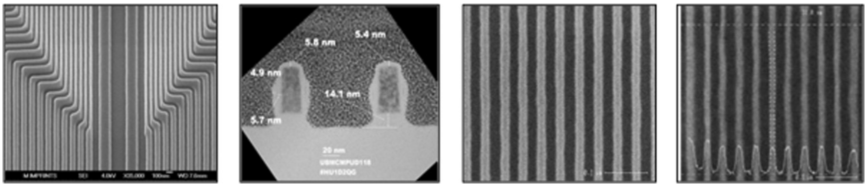


Master imprints in resist



Fused quartz replica mask

9.20 Mask replication process showing the imprints from the master mask and the pattern transferred images on the replica mask.

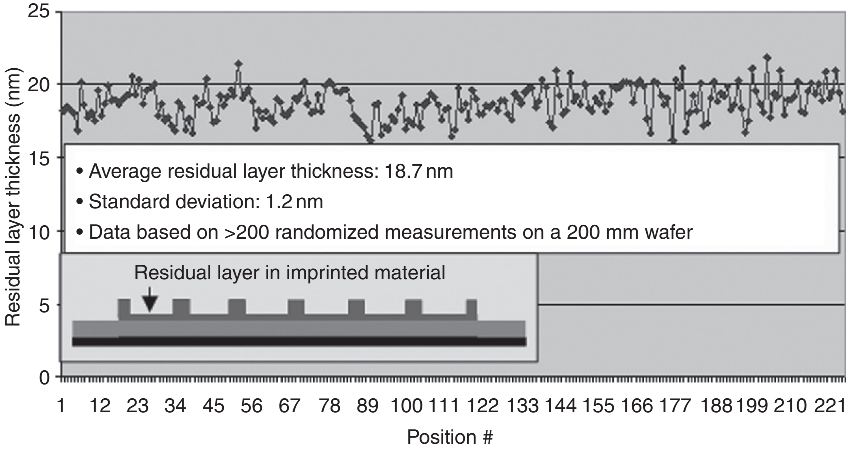


9.21 J-FIL SEMs. From left to right: 38 nm NAND Flash gate layer, 30 nm storage class memory cross section, 22 nm half pitch resist lines and 11 nm half pitch resist lines.

first, it is important to minimize the RLT in order to minimize CD bias during pattern transfer, and secondly it is critical to control the RLT uniformity, in order to avoid any significant CD uniformity resulting from the descum of the residual layer itself. Residual layer control is affected by a variety of factors including:

- *alignment errors between the imprint mask and the substrate*
- field flatness of wafers, imprint masks, and chucks
- Drop-on-demand resist placement relative to the pattern on the mask

In order to minimize CD biases and variations during the removal of the residual layer, a sub-32 nm pattern typically requires a process with a mean residual layer of less than 15 nm. Therefore, in the presence of pattern density variations within a field, resist dispense that is uncorrelated with the mask pattern will require the liquid to travel distances over

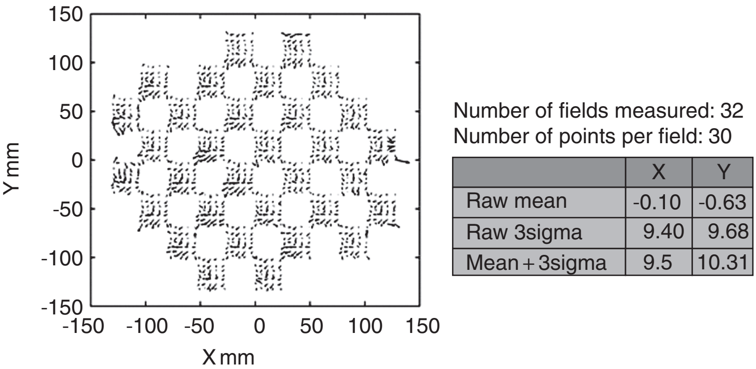


9.22 RLT uniformity across a 200 mm wafer.

millimeters in nanoscale channels. In practice, this situation leads to very high localized pressures in the fluid causing wafer or mask deformations, highly varying residual layers, and longer filling times.[39] In one demonstration, 6 picoliter drops were dispensed in correlation with the mask design, leading to several thousand drops per 26 mm × 32 mm field. This process was automated, and was based on off-line volume computations using mask design information available in GDS-II format. By incorporating the three factors presented above into the tool, a residual layer variation of <5 nm, 3σ, was achieved over a 200 mm wafer, as shown in Fig. 9.22.[40]

Alignment is defined as the accuracy with which an imprinted field can be registered relative to a previously lithographed field at the four corners of the patterned fields. Overlay refers to the accuracy with which an imprinted field can be registered relative to the previous field at the four corners and many points (typically ~100) over a field. The imprint stepper has three key sub-systems that contribute to alignment and overlay:

- ***Interferometric moiré alignment technique (i-MAT):*** This is the approach used to obtain real-time relative overlay errors between points on the imprint mask and the corresponding points on the wafer. The system is capable of measuring alignment errors at a single point well below 1 nm.
- ***Magnification actuator system:*** The magnification actuator works on the basis of imparting elastic deformations to fused silica over a small range of motion (typically on the order of a few ppm). The system incorporates an array of force feedback controlled actuators that are mounted around



Number of fields measured: 32
Number of points per field: 30

	X	Y
Raw mean	-0.10	-0.63
Raw 3sigma	9.40	9.68
Mean + 3sigma	9.5	10.31

9.23 Overlay on a 300 mm wafer using an advanced J-FIL tool. Overlay on the order of 10 nm is achieved across the wafer.

the imprint mask. This strategy allows in-plane corrections in X and Y, orthogonality, and to some extent higher order distortion corrections as discussed in Reference 26.

- ***Wafer stage motion for rigid body corrections (X, Y, and Θ):*** An air bearing wafer stepper stage is used to align the mask and the wafer. The stage needs to hold position with very low noise to maximize overlay performance. The low viscosity liquid resist, in turn, acts as a damping agent for any stage vibrations.

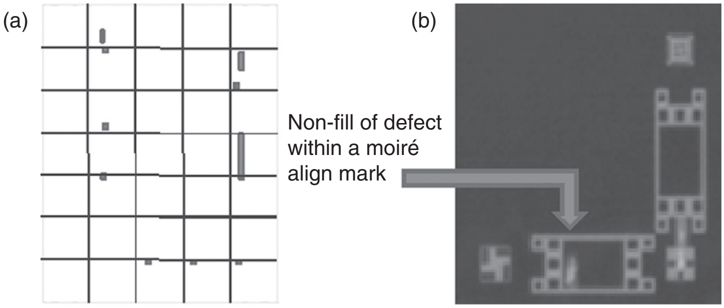
Overlay requirements for a NAND Flash memory device are typically on the order of 1/3 of the half pitch. As an example, for a 32 nm NAND Flash device, a mix-and-match overlay of approximately 10 nm is required. Figure 9.23 depicts a mix-and-match overlay of 10 nm across a 300 mm wafer.[41] Further improvements are likely to be found through thermal management, reduced noise from the magnification actuators, and by reducing the image placement errors on the imprint mask itself.

The cost of any stepper is a function of the process throughput. A throughput of 20 wafers per hour (WPH) is considered a reasonable production target for imprint lithography, and is based on the fact that imprint tools are significantly lower in capital cost as compared to 193 nm immersion and Extreme Ultraviolet (EUV) tools. Based on this target, one shot has to be completed in about 1.6 s (see Table 9.1.).[40]

Of all the steps presented in Table 9.1, only fluid filling is believed to be a fundamental technical risk. The other targets can be met using extensions of known engineering approaches. Fluid filling in a drop-on-demand UV imprint process is affected by several factors, such as fluid viscosity, dispense drop resolution, control of fluid front dynamics, and targeting of drops relative to the mask design. Also, changes in pattern density in the mask design

Table 9.1 Field by field time budget for a throughput of 20 wph (assuming 100 fields/300 mm wafer; and a 20 s overhead for each wafer for loading, set-up and unloading)

Field imprint steps	Time allocation (s)
Stage move, fluid dispense	0.20
Imprint mask down	0.10
Alignment, fluid filling	1.00
UV cure	0.15
Separation	0.15
Total	1.60



9.24 (a) Non-fill defects in a 26 mm × 33 mm field for a resist spread time of only 1.5 s. (b) The non-fill defects were confined to transition regions and to the area near the moiré align mark.

can cause regions where drop tailoring is sub-optimal due to limitations in drop resolution and drop placement.

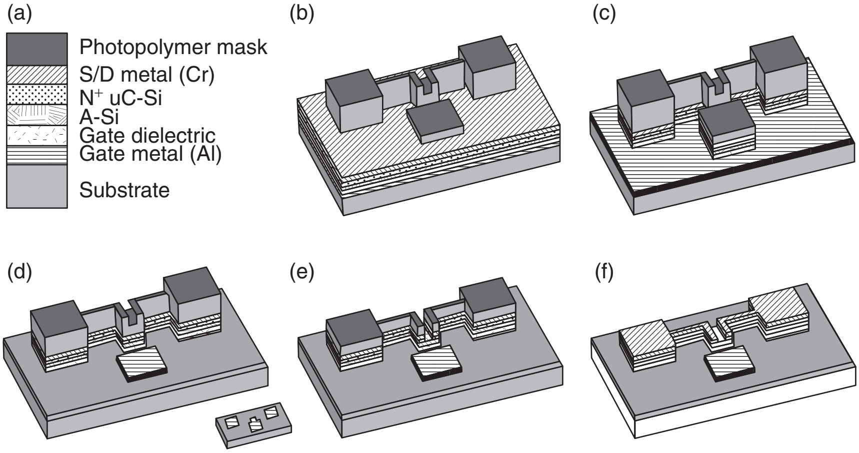
By applying the filling guidelines described above, a test pattern consisting of 24 nm device patterns, dummy fill, and metrology marks was printed and inspected for non-fill defects.[41] In this work, further optimization of the drop pattern was achieved by using gridded drop patterns in areas where the features were essentially parallel lines. In addition, the drop volume was set as low as 1.5 picoliters. The results are shown in Fig. 9.24. Figure 9.24a depicts the location of each defect within the printed mask field for a fill time of 1.5 s. Although the defectivity (1.2/cm^2) was higher than the targeted value of 1.0/cm^2, it should be noted that the defects are systematic. Two types of non-fill defects were observed. The first were the non-fill defects within the printed moiré align mark (see Fig. 9.24b.). The second defect always occurred in transition areas between a repeating structure in a die and another pattern type. Both defects can likely be addressed with specific imprint patterns designed to enhance filling in these areas, and by further reductions in drop volumes.

9.8 Roll to roll imprint lithography

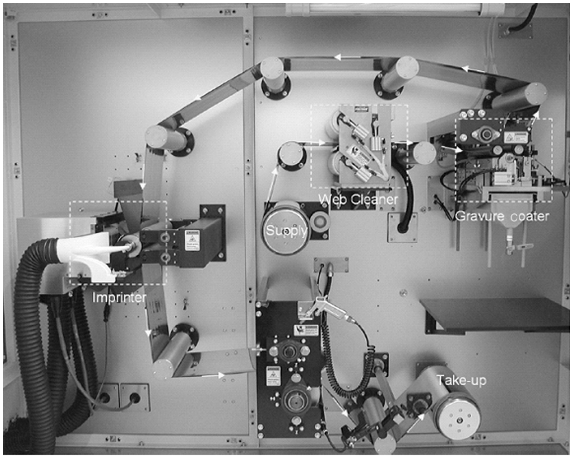
R2R printing, or web printing, involves the patterning of flexible materials such as plastics or metal foils. The flexible material, or web, is unwound from a core, processed, and then returned to a second core at the end of the sequence. Many R2R processes already exist for etch and deposition, and several groups are now developing imprint R2R methods. If successful, the potential applications include thin film transistors, flexible displays, wire grid polarizers, color filters, and solar devices. The challenge, as always, is to create a process that meets both defectivity and throughput requirements.

The Hewlett-Packard Company has recently focused its efforts on flexible displays and has created an active matrix backplane for driving a display using a R2R process and an imprint method referred to as self-aligned imprint lithography (SAIL). The idea of the SAIL process is to create a three dimensional multi-level roller mask and print the 3-D relief image over a film stack consisting of aluminum for the bottom gate metal, a dielectric, an undoped amorphous silicon layer, a microcrystalline silicon layer, and a final layer of chromium, which is patterned to form both the data lines and pixel electrodes.[42,43] A schematic of the SAIL process is shown in Fig. 9.25.

The roller mask is fabricated by first creating a master mask on a planar silicon substrate, and then transferring the pattern to either a PDMS



9.25 Process sequence for HP's SAIL process. (a) Blanket deposited TFT stack; (b) mask on stack with residual etched; (c) device isolated; (d) mask thinned one level, gate metal patterned (detail shows undercut); (e) mask thinned one more level to expose channel; (f) finished SAIL TFT.



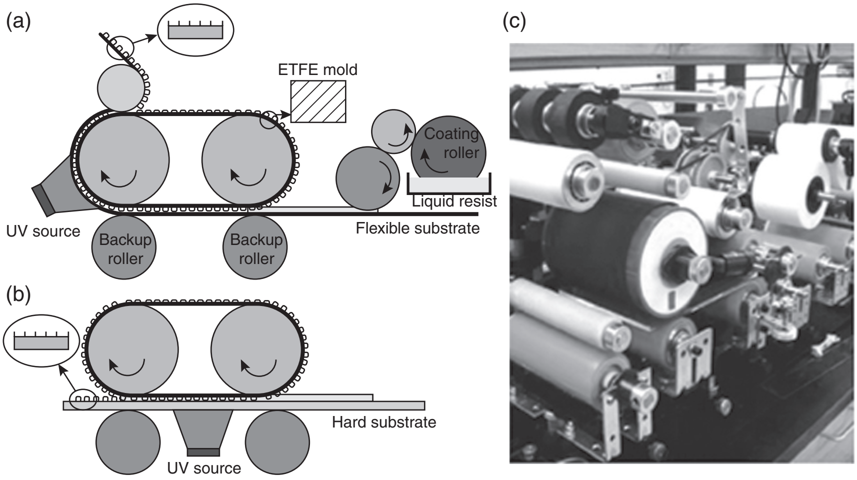
9.26 Web based imprint tool used to create flexible displays.

or TeflonAF flexible film. The film is then peeled away from the silicon, attached to the roller, and used to print the 3-D structures. A photo of the prototype R2R imprint tool is shown in Fig. 9.26. Working color displays consisting of 160 × 120 pixels with a pixel size of 480 μm × 480 μm have been successfully fabricated with this process.

Nanoscale applications include both color filters and wire grid polarizers, and Ahn and Guo have concentrated their efforts on high throughput/resolution R2R systems. Their initial tool focused on simple gratings, and they were able to demonstrate a 100 nm pitch grating using an epoxysilicone resist across a 10 mm width roll.[44] More recently, they have scaled their process and created a large area (6 in.) tool capable of printing a four inch wide film.[45] The system is capable of printing in both an R2R mode and a roll-to-plate (R2P) mode. A schematic and photo of the system is shown in Fig. 9.27. Web speeds of 1m/min were obtained and were limited by the motor used to drive the web.

The epoxysilicone resist was chosen for this study, since the material exhibits a cationic curing mechanism and is not subject to oxygen poisoning typical of acrylate-based resists. In addition, the shrinkage of the resist after curing is substantially smaller than acrylate resists. Typical sensitivity of the epoxysilicone resist is on the order of 100 mW/cm^2. RLT was studied, and the key parameters affecting RLT were web speed and the force applied between the rollers. Not surprisingly, RLT decreases with decreasing web speed and increasing force.

It should be noted that the viscosity of the epoxysilicone resist prohibits the printing of very thin RLTs. Typical values of RLT ranged anywhere from 200 to 300 nm. Clearly, this is a limiting factor for device applications



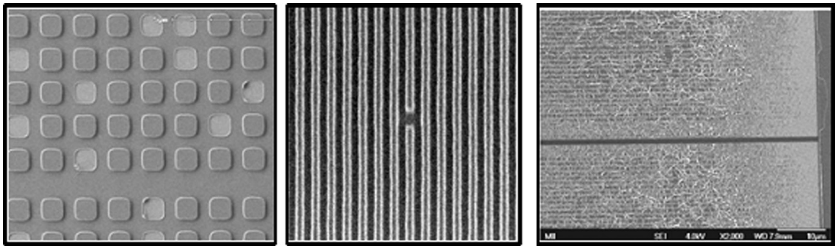
9.27 Schematic of (a) roll to roll and (b) roll-to-plate imprinter. (c) Photograph of the tool. Maximum throughput is 1 m/s.

requiring pattern transfer through the underlying substrate, and low viscosity resist materials and processes will need to be implemented in order to reduce RLTs to less than 20 nm.

9.9 Defectivity

Any lithographic approach needs to satisfy a defectivity specification for a given application. Semiconductor applications typically strive for defect levels less than 0.10 defect/cm^2. For the hard drive industry, the final media is expected to have an aerial defect ratio of less than 10^{-4}.

Early work on step-and-flash imprint lithography indicated that the low surface energy monolayer applied to the template acted effectively as a self-cleaning agent. This attribute has been reported in several publications.[20,27] While the data clearly illustrate a self-cleaning effect, this does not provide a sufficient criterion for taking the technology into manufacturing. Defectivity of the overall process must be considered, and defect sources must be broken down and analyzed into their various components. The mask, the resist, and the imprint tool all must work hand in hand to create a process that is sustainable for many thousands of imprints. As an example, several papers have been published on the use of 1H, 1H, 2H, 2H-FOTS anti-adhesive monolayer. The FOTS is successful at lowering surface energy and enabling separation of the mask and substrate. It is well known, however, that the monolayer has a limited lifetime, and that defects are readily observed after only tens of imprints.[8]

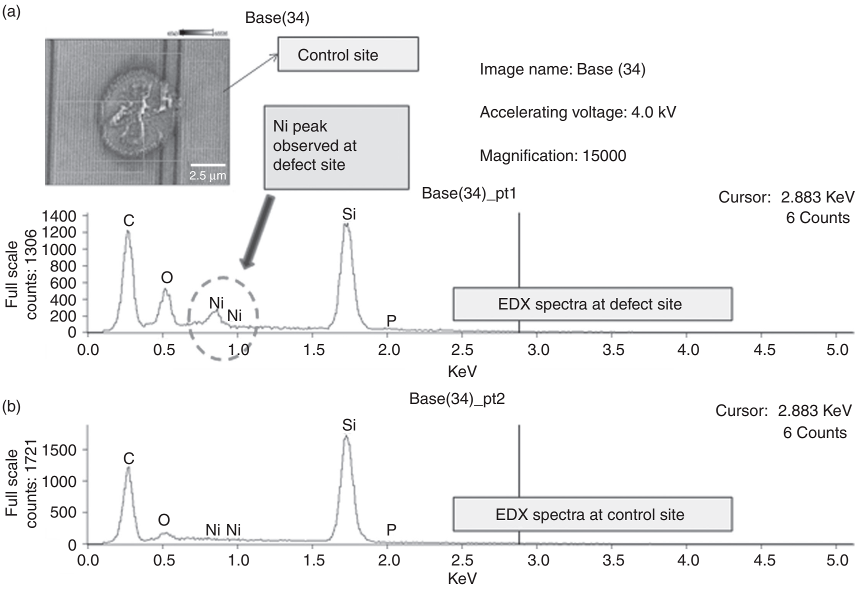


9.28 From left to right: non-fill defect, plug defect, and a transition region defect.

The imprint process always starts with the fabrication of a mask. Defects created in the mask during the fabrication process must be inspected and repaired as required to meet defect tolerances. Cleaning of the mask is critical to remove particulates and contamination on the surface that may impede filling. The resist must be designed to adhere to the substrate surface, release from the mask surface, and be filtered so that the material itself is not a defect source. The imprint process must result in fast filling and avoid non-fill defects (as discussed in Section 9.7). In addition, the process, including filling and separation, must be robust and immune to defects such as line collapse and feature shearing. Finally, and most importantly, particulates must be well controlled, as they have the potential to damage the mask, thereby creating a permanent defect in the mask. Some defect examples are shown in Fig. 9.28.

Defects are generally placed into two categories: random and repeating. Random defects include non-fill defects, and are typically controlled through process optimization. Imprint defects caused by particles can become repeater defects and are therefore the key to minimizing overall defectivity. In a recent study by Singh *et al.*, repeater defectivity was studied by printing 2000 fields using a 120 nm NAND Flash-like layout.[41] To eliminate wafer related particles, only fields that were confirmed to have particles smaller than 70 nm (as measured using a KLA-T SP1 blank wafer inspection system) were printed. The patterned wafers were inspected using the KLA-T 2132 and the repeater defects locations were identified. Composition of the defect was then determined by placing the wafer in an SEM, driving to the location where the repeater defect was first reported, and performing an Energy-dispersive X-ray spectroscopy (EDX) measurement at the defect site.

All but one of the defects were confirmed to contain nickel. Figure 9.29 shows both a picture of the particle and the EDX measurements. Figure 9.29a shows the EDX signal within the defective area. Figure 9.29b is a scan of a clean pattern. The nickel peak was clearly observed, and every defect containing nickel showed a comparable signal intensity. It was noted that both



9.29 Results for an extended imprint run. The particles that created printing defects (see for example the defect in the center of the SEM image just above the scan in Figure 9.29a) were found to contain nickel.

the inkjet dispenser and the filter unit contained within the imprint tool contain nickel parts.

In a similar study, using more state-of-the-art processing tools, Malloy and Lloyd were able to demonstrate defect densities as low as 0.09 defects/cm^2 across a 300 mm silicon wafer.[46] The next step is to confirm low defectivity for long runs consisting of as many as 50 000 imprints. Passage through this gate is critical to introducing the technology into high volume manufacturing.

9.10 Conclusions

Imprint lithography has come a long way in a very short period of time. Resolution seems limited to the ability to form a relief image in the template and sub-10 nm printing has already been demonstrated. The technology is cost effective, and applications will continue to grow as emerging markets in the field of nanotechnology continue to blossom.

To be considered a method for fabricating silicon integrated circuits, several concerns including overlay, throughput and, most critically, defectivity need to be addressed. UV-NIL and, in particular, J-FIL seems the best imprinting option for meeting the stringent requirements of future generations of silicon-based circuitry. Tools, templates, and resists are readily available to start exercising the technology and will be used to answer the open issues. If these issues can be solved, imprint lithography may be the right

NGL, since extendibility beyond 10 nm has already been demonstrated. The last consideration then becomes the supporting infrastructure. Projection lithography has been in the mainstream now for over 30 years, and the ability to write, inspect, and correct a 1X template will need to be developed. Electron-based inspection and repair tools, as well as faster high resolution electron beam writing systems, may provide the pathway for template fabrication in the very near future.

With respect to emerging markets, the application of imprint lithography is enabling. No other lithographic technology is capable of meeting the stringent cost of ownership demands necessary to bring products to market. Patterned media is already one promising application. It will be interesting to see where imprint technology is eventually applied to address the power, energy and biomedical markets.

9.11 Acknowledgments

The writing of this chapter would have been impossible without the contributions of so many different and talented scientists. I would like to thank several collaborators in particular. David Mancini, Bill Dauksher, and Kevin Nordquist are responsible for much of the early work on S-FIL templates. C. Grant Willson and S. V. Sreenivasan provided most of the expertise for the tool and resist discussions. Se Hyun Ahn provided the expertise on R2R imprinting. Finally, I would like to thank all of the engineers and scientists at Molecular Imprints. These are the folks who design and build the imprint tools and solve the problems critical to bringing the technology to production.

9.12 References

1. D. J. Resnick, W. J. Dauksher, D. Mancini, K. J. Nordquist, T. C. Bailey, S. Johnson, N. Stacey, J. G. Ekerdt, C. G. Willson, S. V. Sreenivasan and N. Schumaker, *Proc. SPIE,* **5037**, 12 (2003).
2. M. Melliar-Smith, S.V. Sreenivasan and D. J. Resnick, 'Step and Flash Imprint Lithography for high volume manufacturing', presented at the 2010 Litho Forum, (2 December 2010).
3. M. Melliar-Smith, presented at the panel discussion on Economics of Lithography for Alternative Applications', SPIE Advanced Lithography Symposium, (2 March 2011).
4. F. Hua, Y. Sun, A. Gaur, M. A. Meitl, L. Bilhaut, L. Rotkina, J. Wang, P. Geil, M. Shim, J. A. Rogers* and A. Shim, *Nano Lett.*, **4**(12), 2467–2471 (2004).
5. Y. Xia and G. M. Whitesides, *Angew. Chem. Int.*, **37**, 550–575 (1998).
6. J. A Rogers, K. E. Paul and G. M. Whitesides, *J. Vac. Sci. Technol. B* **16**(1), 88 (1998).
7. S. Y. Chou, P. R. Krauss and P. J. Renstrom, *J. Vac. Sci. Technol. B*, **14**(6), 4129–4133 (1996).

8. M. M. Alkaisi and K. Mohamed, 'Lithography', edited by Michael Wang, ISBN 978–953–307–064–3, pp. 656, (February 2010).
9. S. Y. Chou, P. R. Krauss, W. Zhang, L. Guo and L. Zhuang, *J. Vac. Sci. Technol. B* **15**(6), 2897–2904 (1997).
10. Z. Yu, S. J. Schablitsky and S. Y. Chou, *Appl. Phys. Lett.*, Vol. **74**, No. 16, 2381–2383 (1999).
11. Y. Chen, D. Macintyre, E. Boyd, D. Moran, I. Thayne and S. Thoms, *J. Vac. Sci. Techno. B* **20**(6), 2887–2890 (2002).
12. X. Cheng, Y. Hong, J. Kanicki and L. Jay Guo, *J. Vac. Sci. Technol. B* **20**(6), 2877–2880 (2002).
13. W. Wu, B. Cui, X. Sun, W. Zhang, L. Zhuang, L. Kong and S. Y. Chou, *J. Vac. Sci. Technol. B* **16**(6), 3825–3829 (1998).
14. Z. Yu, W. Wu, L. Chen and S. Y. Chou, *J. Vac. Sci. Technol. B*. **19**(6), 2816–2819 (2001).
15. B. E. Little and S. T. Chu, *Opt. Photonics News* **11**,24 (2000).
16. C. Chao and L. Jay Guo, *J. Vac. Sci. Technol. B* **20**(6), 2862–2866 (2002).
17. C. M. Falco, J. M. van Delft, J. P. Weterings, A. K. van Langen-Suurling and H. Romijn, *J. Vac. Sci. Technol. B* **18**(6), (Nov/Dec 2000).
18. S. Y. Chou, C, Keimel and J, Gu, *Lett. Nat.*, **417**, 835–837 (2002).
19. M. Otto, M. Bender, B. Hadam, B. Spangenberg and H. Kurz, *Microelectron. Eng.,* **57–58,** 361–366 (2001).
20. M. Colburn, S. Johnson, M. Stewart, S. Damle, T. Bailey, B. Choi, M. Wedlake, T. Michaelson, S. V. Sreenivasan, J. Ekerdt and C. G. Willson, *Proc. SPIE*, *Emerging Lithographic Technologies* III, **379** (1999).
21. M. Colburn, I. Suez, B. J. Choi, M. Meissl, T. Bailey, S. V. Sreenivasan, J. G. Ekerdt and C. G. Willson, *J. Vac. Sci. Technol. B,* **19**(6), 2685 (2001).
22. F. Xu, N. Stacey, M. Watts, Van Truskett, I. McMackin, J. Choi, P. Schumaker, E. Thompson, D. Babbs, S.V. Sreenivasan, G. Willson and N. Schumaker, 'Development of Imprint Materials for the Step and Flash Imprint Lithography Process,' presented at the SPIE Proceedings on Advanced Lithography, Emerging Lithography Technologies, Santa Clara, CA, (March 2004).
23. B. J. Choi, S. Johnson, M. Colburn, S.V. Sreenivasan and C. G. Willson, *J. Int. Soc. Precis. Eng. Nanotechnol.*, **25**, 3, 192–199 (July 2001).
24. E. E. Moon, J. Lee, P. Everett and H. I. Smith, *J. Vac. Sci. Technol. B*: *Microelectr. Nanometer Struct.*, **16**(6), 3631 (November 1998).
25. B. J. Choi K. J. Nordquist, A. Cherala, L. Casoose, K. Gehoski, W. Dauksher, S.V. Sreenivasan and D. J. Resnick, 'Distortion and overlay performance of UV step and repeat imprint lithography,' *Microelectron. Eng.*, 2004 **78–79**, 633, (January 2005).
26. A. Cherala, B. J. Choi, P. K. Nimmakayala, M. J. Meissl and S. V. Sreenivasan, 'An Apparatus for Varying the Dimensions of a Substrate During Nano-Scale Manufacturing,' US Patent No. 7,170,589 (2007).
27. M. Colburn, T. Bailey, B. J. Choi, J. G. Ekerdt and S. V. Sreenivasan, *Solid State Technol.*, **67** (June 2001).
28. K. H. Smith, J. R. Wasson, P. J. S. Mangat, W. J. Dauksher and D. J. Resnick, *J. Vac. Sci. Technol. B* **19**(6), 2906 (2001).
29. T. C. Bailey, D. J. Resnick, D. Mancini, K. J. Nordquist, W. J. Dauksher, E. Ainley, A. Talin, K. Gehoski, J. H. Baker,, B. J. Choi, S. Johnson, M. Colburn, S. V.

Sreenivasan, J. G. Ekerdt and C. G. Willson, *Microelectron. Eng.,* **61–62,** 461–467 (2002).

30. W. J. Dauksher, K. J. Nordquist, D. Mancini, D. J. Resnick, J. H. Baker, A. E. Hooper, A. A. Talin, T. C. Bailey, A. M. Lemonds, S. V. Sreenivasan, J. G. Ekerdt and C. G. Willson, *J. Vac. Sci. Technol. B,* **20**(6), 2857–2861 (2002).
31. D. J. Resnick, W. J. Dauksher, D. Mancini, K. J. Nordquist, E. Ainley, K. Gehoski, J. H. Baker, T. C. Bailey, B. J. Choi, S. Johnson, S. V. Sreenivasan, J. G. Ekerdt and C. G. Willson, *Proc. SPIE,* **4688**, 205 (2002).
32. N. Hayashi, T. Abe, T. Shimomura, Y. Inazuki, T. Takikawa and H. Mohri, *Proc. SPIE,* **7985**, 798505 (2011).
33. L. J. Myron, E. Thompson, I. McMackin, D. J. Resnick, T. Kitamura, T. Hasebe, S. Nakazawa, T. Tokumoto, E. Ainley, K. Nordquist and W. J. Dauksher, *Proc. SPIE,* **6151**, 61510M (2006).
34. K. Selinidis, E. Thompson, S. V. Sreenivasan and D. J. Resnick, *Proc. SPIE,* **7470**, 74700I (2009).
35. M. Pritschow, H. Dobberstein, K. Edinger, M. Irmscher, D. J. Resnick, K. Selinidis, E. Thompson and M. Waiblinger, *Proc. SPIE,* **7488**, 74880V (2009).
36. C. Brooks, G. M. Schmid, M. Miller, S. Johnson, N. Khusnatdinov, D. LaBrake, D. J. Resnick and S. V. Sreenivasan, *Proc. SPIE,* **7271**, 72711L (2009).
37. C. Brooks, K. Selinidis, G. Doyle, L. Brown, D. LaBrake, D. J. Resnick and S. V. Sreenivasan, *Proc. SPIE,* **7823**, 78230O (2010).
38. S. Kosta Selinidis, Cynthia B. Brooks, Gary F. Doyle, L. Brown, C. Jones, J. Imhof, L. Dwayne LaBrake, J. Douglas Resnick and S.V. Sreenivasan, *Proc. SPIE,* **7970**, 797009 (2011).
39. V. Sirotkin, A. Svintsov, H. Schift and S. Zaitsev, 'Coarse-grain method for modeling of stamp and substrate deformation in nanoimprint,' *Microelectron. Eng.*, **84**, 868 (2007).
40. S. V. Sreenivasan and P. D. Schumaker, 'Critical dimension control, overlay, and throughput budgets in uv nanoimprint stepper technology', B.J. Choi, ASPE, Spring Proceedings, 7–8 April, Berkley, California, (2008).
41. L. Singh, K. Luo, Z. Ye, F. Xu, G. Haase, D. Curran, D. LaBrake, D. Resnick and S. V. Sreenivasan, *Proc. SPIE,* **7970**, 797007 (2010).
42. A. Jeans, M. Almanza-Workman, R. Cobene, R. Elder, R. Garcia, F. Gomez-Pancorbo, W. Jackson, M. Jam, H. Kim, O. Kwon, H. Luo, J. Maltabes, P. Mei, C. Perlov, M. Smith, C. Taussig, F. Jeffrey, S. Braymen, J. Hauschildt, K. Junge, D. Larson and D. Stieler, *Proc. SPIE,* **7637** 763719–763712 (2010).
43. E. R. Holland, A. Jeans, P. Mei, C. P. Taussig, R. E. Elder, C. Bell, E. Howard and J. Stowell, *Proc. SPIE,* **7970**, 797016 (2011).
44. S. H. Ahn and L. J. Guo, *Adv. Mater,*.2044–2049 (2008).
45. S. H. Ahn and L. J. Guo, *ACS Nano,* **3**(8), 2304–2310 (2009).
46. M. Malloy, L. C. Litt, S. Johnson, D. J. Resnick and D. Lovell, *Proc. SPIE,* **7970**, 797006 (2011).

10

Nanostructures: fabrication and applications

X. CHENG, Texas A&M University, USA

DOI: 10.1533/9780857098757.348

Abstract: Nanostructures refer to materials or structures that have at least one dimension between 1 and 100 nm. The various sub-fields of nanotechnology focus on the creation, characterization, and application of nanostructures. In this chapter, we discuss the fabrication and application of nanostructures commonly studied and used in nanotechnology. Nanostructures are fabricated by methods as simple as solution processing, and as complicated as advanced lithography. In addition to extensive efforts in the fabrication of nanostructures, current research and development of nanostructures also focus on the investigation of fundamental material properties, including mechanical, electrical, and optical properties at the nanoscale, and the development of novel applications in a broad range of engineering fields. In this chapter, we review some unique properties of nanostructures and discuss the most actively pursued applications.

Key words: nanostructure, quantum dot, quantum well, lithography, plasmonic.

10.1 Introduction

Nanostructures are materials with nanometer scale sizes in one, two, or three dimensions. The term 'nanostructure' is loosely defined, but typically the dimensions of nanostructures are between 1 and 100 nm. Nanostructures in recent years have generated a great deal of interest, from basic scientific research to commercial development. Nanostructures bring forth novel material properties and superior performance in certain areas, which has sparked much enthusiasm in this emerging area. Today the study of nanostructures, broadly known as nanotechnology, has evolved into a truly multidisciplinary effort made by experts from almost every traditional science and engineering field.

In general, nanostructures can be divided into surface and bulk nanostructures. Surface nanostructures are nanoscale patterns created on the surface of a substrate. The most prominent example is the nanoscale electronic structures fabricated on silicon wafer surface using advanced microelectronic fabrication techniques. Bulk nanostructures refer to individual nanomaterials, or an assembly of nanomaterials. For example, colloidal nanostructures are nanoparticles or nanowires dispersed in a liquid with the help of surfactants, and nanocomposites are ordered or disordered distributions of nanomaterials in another phase.

The most commonly studied nanomaterials include carbon-based nanostructures, quantum dots, metallic nanomaterials, and biological nanostructures. Among all nanostructures, carbon nanotubes (CNT) and graphene are probably the best known and most widely studied. The widespread interest in carbon-based nanostructures originates from their superior, sometimes unique, physical properties, providing the prospect of breakthroughs in many important engineering applications (Avouris *et al.*, 2007). Nanostructures based on traditional compound semiconductors, such as quantum dots and quantum wells, have been well developed during the last three decades. The quantum confinement in those nanostructures provides superior optoelectronic properties, such as electronic structure engineering and density of states tuning. The maturity of those nanostructures has led to many important commercial applications such as LEDs, solid-state lasers and photodetectors. Plasmonic excitation in metallic nanostructures has also generated a lot of excitement for its ability to strongly focus electromagnetic radiation at the infrared and visible wavelengths. A prominent example is surface-enhanced Raman spectroscopy, where metallic nanostructures greatly enhance the usually-weak Raman scattering signal for exceptionally sensitive molecule detection and identification. The plasmonic modes on metallic nanostructures have extremely small spatial distribution, enabling them for high-resolution imaging and lithography. Nanostructures also come with novel thermal and magnetic properties, which are gradually being exploited in certain applications. In this chapter, we attempt to provide only a broad description of the most common aspects of nanostructure fabrication and application. More detailed information on each topic can be found in many review papers and book chapters cited in this chapter.

It is worth mentioning that, despite the definition, there are structures and materials that have nanoscale dimensions that are not usually discussed in the context of nanotechnology. For example, individual linear polymer chains can be regarded as one-dimensional 'nanowires,' but the study of polymers is usually not regarded as nanotechnology, unless the polymers are mixed with other nanomaterials to form nanocomposites. In the microelectronic fabrication process, thin film deposition is an important and widely used technique. The metallic and dielectric thin films can have a thickness in the range of a few nanometers to 100 nm. However, those thin films are used for their bulk properties (e.g., electrodes for current conduction or gate dielectric for the insulation of carrier transport), not for the unique properties brought forth by their nanoscale dimensions. Similar examples can be found in organic electronics, where most of the layers used in organic light-emitting device and organic thin film transistors typically have a thickness of less than 100 nm.

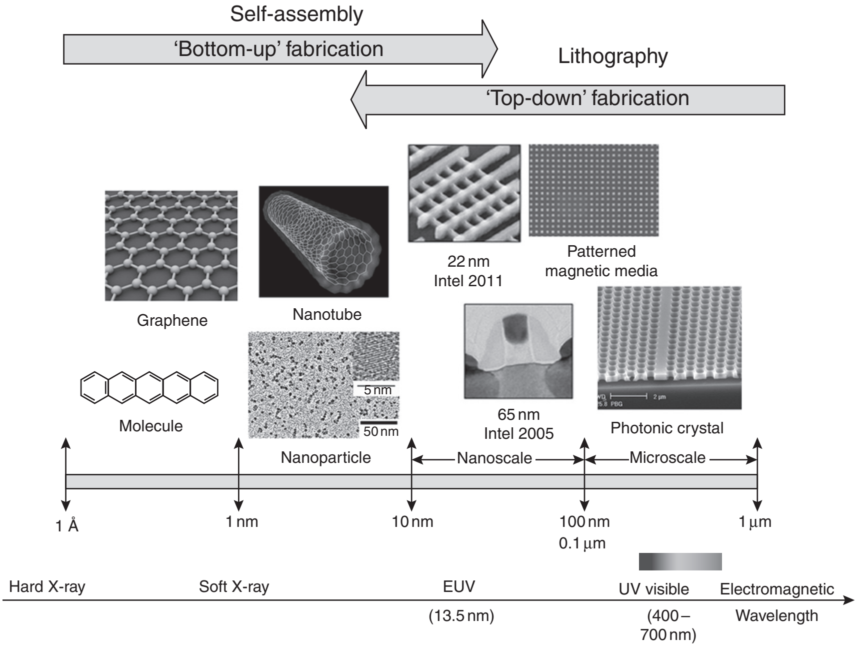
10.2 Characterization of nanostructures

Due to their extremely small dimensions, the characterization of nanostructures usually requires dedicated tools and special techniques. The most common characterization of nanostructures is in terms of shape and morphology. This information is usually obtained by various imaging techniques. Most high-resolution imaging techniques share the same principles as high-resolution lithography techniques, because these two techniques complement each other. In high-resolution imaging, tightly focused photon or electron beams scan the nanostructures to construct an image. Techniques such as scanning electron microscope (SEM), near-field scanning optical microscope (NSOM), and plasmonic imaging are based on this principle. If a sharp mechanical tip is used, high-resolution imaging can also be obtained, such as the atomic-force microscope (AFM), the scanning-tunneling microscope (STM), and their variations. Electron transmission can also be used to probe and image nanostructures. This feature has been adopted in the transmission electron microscope (TEM) to achieve atomic-level resolution when a high-energy electron-beam shoots through a thin layer of nanostructures.

In the characterization of individual nanostructures, often a dedicated device structure needs to be fabricated to interface between the nanostructure and the macroscopic measurement system. For example, electrical characterization of individual nanoparticles or nanowires requires the patterning of contact electrodes. This is very challenging because of the extremely small size of the nanostructures. Other factors, such as very low signal level, can add to measurement complexity. Highly sensitive electronic systems and sophisticated environmental chambers are usually needed in the probe station for nanostructure characterization.

10.3 Methods to create nanostructures: top-down fabrication of nanostructures

Nanostructures can be generated by a wide variety of techniques. In general, surface nanostructures are often defined and fabricated by techniques developed in microelectronic processing. Those methods start from a large (in relation to the nanostructures to be created) substrate. Nanoscale patterns are defined by advanced lithography techniques, followed by material removal (etching) or material deposition techniques. This approach is generally known as the ‘top-down’ fabrication of nanostructures. Alternatively,


Self-assembly
'Bottom-up' fabrication
Lithography
'Top-down' fabrication
Graphene
Nanotube
22 nm
Intel 2011
Patterned
magnetic media
Molecule
5 nm
50 nm
Nanoparticle
65 nm
Intel 2005
Photonic crystal
Nanoscale
Microscale
1 Å
1 nm
10 nm
100 nm
0.1 µm
1 µm
Hard X-ray
Soft X-ray
EUV
(13.5 nm)
UV visible
(400–700 nm)
Electromagnetic
Wavelength

10.1 The 'top-down' and 'bottom-up' routes to create nanostructures and their characteristic lengths scales.

nanostructured materials are often created by chemical synthesis. The assembly of atoms or molecules, with the help of autonomous or carefully designed size limiting mechanisms, can result in materials with one, two, or three dimensions at the nanoscale. Those synthesized nanomaterials can then be further assembled into larger structures, such as super-lattices. This approach of building nanostructures and nanomaterials from atoms and molecules is generally referred to as the 'bottom-up' fabrication of nanostructures. As shown in Fig. 10.1, nanostructures are becoming the converging spot for 'top-down' and 'bottom-up' approaches, thus bringing scientists and engineers from all disciplines into the exciting field of nanotechnology.

The process of creating pre-designed patterns on a surface is called lithography. Lithography has evolved with the microelectronic industry. The demand for fabrication of ever-smaller electronic devices has stimulated the development of many advanced lithography techniques to fabricate nanoscale structures.

In terms of their working principles and operating schemes, nanoscale lithography techniques can be categorized into energy-beam-based lithography, tip-based lithography, and mechanical forming. In terms of their application objectives, those lithography techniques can be divided into two types:

initial pattern generation and pattern replication. Initial pattern generation at the nanoscale usually requires sophisticated and expensive equipment. It is usually featured by very high resolution, but high cost and low throughput are the main issues associated with the initial pattern generation. To meet the low cost and high throughput demand for practical applications, pattern replication techniques are indispensable. Pattern replication techniques use master copies produced from initial pattern generation techniques, and produce duplicate copies with low cost and high throughput.

10.3.1 Energy-beam-based lithography: photolithography, electron-beam lithography, and holographic lithography

Energy-beam-based lithography techniques exploit the interaction of an energy beam, such as a photon, electron, or ion-beam, with a thin film active material called a resist. In energy-beam-based lithography techniques, the energy beam is focused into a very tight spot or projected onto a resist layer. The interaction between the energy beam and the resist layer changes the properties of the resist material. Pattern can be generated by scanning the energy beam over the resist surface, like drawing a graph with a pencil. The different properties in exposed and unexposed areas in the resist layer result in selective resist removal in a subsequent 'developing' step.

Photolithography is currently the mainstream technology used in modern microelectronic fabrication. The objective of photolithography is to create light intensity pattern in light-sensitive photoresist thin film on wafer surface. The light intensity pattern is defined by the photomask, which contains the layout structures to be 'printed' on the surface of a wafer. In general, there are two modes of photolithography: projection mode and contact mode. In projection mode (Chapter 1), the image on the photomask is projected onto the surface of photoresist. This process is in fact equivalent to a camera action recording the image of the photomask on a recording medium, which is the photoresist. The projection photolithography tool can be viewed as a 'camera' that uses UV wavelength as a light source and photoresist as an image recording medium; thus, many features of projection lithography can be understood using the concepts of traditional photography.

The resolution of photolithography is described by a simple equation: $l = k_1\lambda/\mathrm{NA}$, where k_1 is the process parameter whose value is determined by the properties of the photoresist and the photoresist processing parameters, λ is the wavelength of the light source, and NA is the numerical aperture of the projection lens. For sub-100 nm nanoscale patterning, it requires the projection tool to use short wavelengths. Currently, 193 nm deep-UV light is used in popular photolithography systems. Active research is going on to further reduce this wavelength into extreme UV, even to X-ray range, in spite of the tremendous technical challenges. In addition to reducing exposure wavelength, increasing the NA can also improve pattern resolution. A simple technique to increase the NA is by immersing the projection lens in a liquid with a higher refractive index than air. Those techniques have enabled the 193 nm immersion photolithography to be widely used in sub-100 nm microelectronic patterning. In addition to decreasing the exposure wavelength and increasing the NA, many other techniques have been developed for photolithography to further enhance pattern resolution. A few well-known examples are double patterning, double exposure, and spacer double patterning (Willson and Roman, 2008).

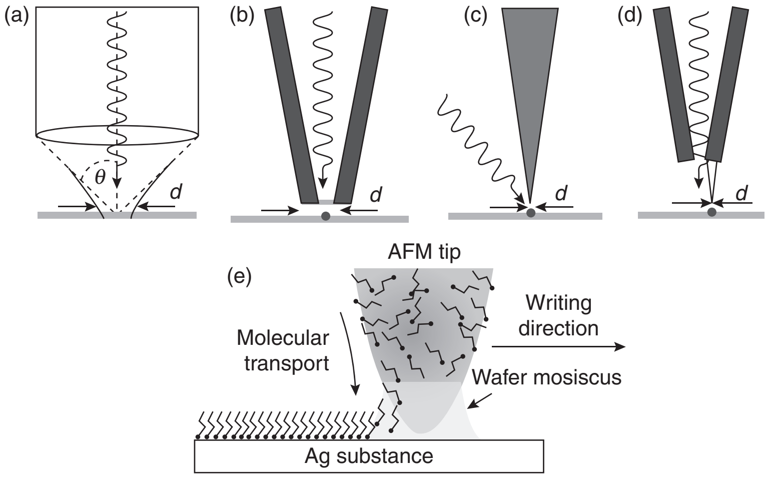
The second mode of photolithography is contact mode, whereby the shadows of photomask patterns are directly cast on a photoresist layer that is in contact with, or in close proximity to, the photomask. This technique is very simple and low cost. The issue is the contact between the photomask and photoresist, which causes defects, and potentially contamination and damage, to the photomask. Contact mode photolithography is primarily used in research or a small-scale fabrication environment.

Photolithography is a pattern replication process. The master copy used in photolithography, the photomask, has to be created by high-resolution pattern generation techniques (Chapter 5). Currently, the most common techniques are laser direct writing and electron- or ion-beam lithography. Electron- and ion-beam lithography techniques use electrooptics to focus a beam of electrons or ions into a small spot, then scanning the resist layer by directing the beam with electronic deflectors and shutters. High acceleration voltage enables the charged beams to be focused down to 1 or 2 nm, which is ideal for very high-resolution pattern generation. However, the serial scanning in charge beam lithography techniques seriously limits their throughput. Thus, those techniques are mainly used in the fabrication of master copies, such as photomasks and nanoimprint templates, that are used in other high throughput pattern replication techniques. There are also several attempts to develop parallel writing or projection electron-beam techniques, although great technical challenges exist.

Another important patterning technique for creating large area periodic structures is holographic lithography (Schattenberg *et al.*, 1990; Brueck 2005), sometimes called interference lithography. The interference of two or multiple coherent optical beams creates a periodic spatial modulation of light intensity, which can be imaged by a resist layer to generate periodic structures. Due to the nature of interference, subwavelength features can be generated over a large area in a single exposure. The high-resolution, high-speed and large area features are attractive in the fabrication of periodic nanostructures.

10.3.2 Tip-based lithography – near-field tip lithography and dip-pen lithography

Commercial photolithography and electron-beam lithography systems are complicated and expensive. To enable low cost patterning of nanostructures, many tip-based lithography techniques have been developed (Sun *et al.*, 2002). Instead of using a tightly focused energy beam, a sharp mechanical tip is used for nanofabrication .The pattern formation mechanisms can vary greatly with different tip-based lithography techniques, as shown in Fig. 10.2. Tapered optical fibers have the capability to focus light into a very small area. When the tip surface is coated with noble metal such as gold and silver, the excitation of surface plasmon (SP) polaritons can lead to exceptionally high intensity and close spatial confinement, as discussed below in the plasmonic section. If the substrate is placed within the near-field of the tip, the tight optical energy spot can achieve pattern writing with deep subwavelength resolution. The biggest advantage of tip-based lithography techniques is their simple construction and low cost, as compared to much more complicated and expensive photolithography and electron-beam lithography systems. Thus, the tip-based lithography system is ideal for writing nanostructures for basic research or prototype development in research labs.



10.2 Schematics of tip-based imaging and lithography system: (a) traditional far-field lens for imaging and projection photolithography (b) NSOM tip (c) tip-enhanced imaging and lithography based on the excitation of plasmonics (d) tip-on-aperture (TOA) structure for plasmonic lithography. (*Source:* Reprinted with permission from (Hartschuh 2008). Copyright: Wiley-VCH Verlag GmbH & Co. (2008).) (e) dip-pen lithography for chemical and biological pattern writing. (*Source:* Reprinted with permission from (Piner, Zhu *et al.* 1999). Copyright: AAAS (1999).)

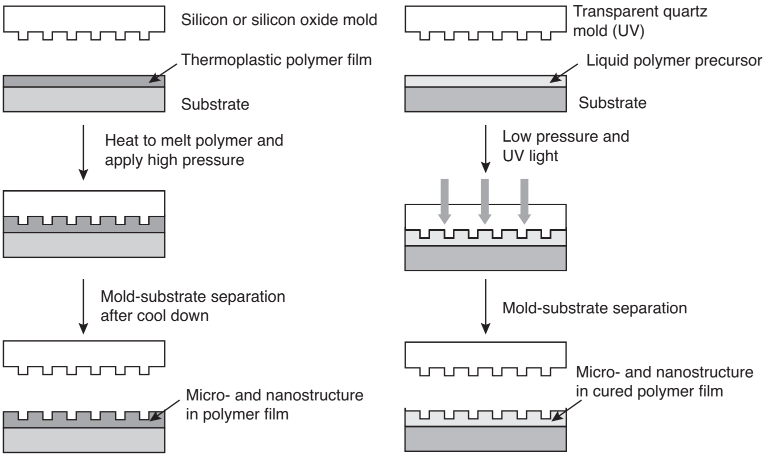
In many cases, patterning chemical and biological entities are desired for chemical, biological, and biomedical engineering applications. Although such functional patterns can be generated by multiple steps of lithography and pattern transfer, the chemicals and energy radiation are often incompatible with the desired functional patterns. Dip-pen lithography has been developed to address this issue (Piner *et al*., 1999; Ginger *et al*., 2004). Instead of using optical energy, the tip in dip-pen lithography can be soaked in a chemical 'ink' and then the 'ink' can be transferred to a substrate surface, much like writing on a paper using a quill pen. The ink is formulated to contain the functional chemicals, or biological molecules such as oligonucleotides and protein to be patterned (Salaita *et al*., 2007). Dip-pen lithography allows quick patterning of nanoscale chemical and biological patterns without the characteristically tedious and potentially detrimental fabrication steps.

In all tip-based lithography, there are two major challenges. The first is the low writing throughput. Since the pattern writing is done using a very sharp tip, and the writing is done in serial scanning, the surface area that can be written is very limited. Although electron-beam lithography also uses a sharp electron beam, the beam scanning is done electronically. In tip-based lithography, the tip scanning is achieved with mechanical actuation, which is inherently much slower than electronic beam scanning. To address this limitation, current research aims at developing a tip-based lithography system with massive parallel tips (Hong and Mirkin, 2000). A large area pattern can be divided into many sub-areas, each of which would be written by a tip.

The second major challenge in tip-based lithography is the need to maintain the distance between the tip and the substrate. In scanning probe lithography, the substrate should be within the near-field distance of the tip, which is usually less than the wavelength of the light. In dip-pen lithography, the tip is expected to make mechanical contact with the substrate. Maintaining accurate tip-substrate separation is important not only for high-resolution pattern writing, but also for protecting the fragile tip. Considering the non-flatness of a large substrate and potential pre-existing surface patterns, it is imperative to include a live feedback mechanism to maintain tip-substrate distance in the tip-based lithography system. In dip-pen lithography, using more flexible materials such as PDMS, to fabricate the tip (Huo *et al*., 2008) or to add a PDMS cushion layer between the tip and its base, can mitigate the contact damage to both the tip and the substrate surface(Shim *et al*., 2011).

10.3.3 Mechanical forming: nanoimprint lithography

Although it is critically important to have the capability to fabricate nanostructures in the first place, it is equally important that the nanostructures are replicated with low cost and high throughput for nanotechnology to


Silicon or silicon oxide mold
Thermoplastic polymer film
Substrate
Heat to melt polymer and apply high pressure
Mold-substrate separation after cool down
Micro- and nanostructure in polymer film
Transparent quartz mold (UV)
Liquid polymer precursor
Substrate
Low pressure and UV light
Mold-substrate separation
Micro- and nanostructure in cured polymer film

10.3 Schematics of NIL. Left: thermal nanoimprint; Right: UV-curable nanoimprint.

make a significant impact on society. A great example of the combination of the initial pattern generation and the pattern replication is electron-beam lithography and photolithography. Electron-beam lithography is indispensable for creating the high-resolution and high-accuracy master copy, the photomask, to be used by photolithography to reproduce the nanostructures in large scale for commercial microelectronic fabrication.

As the feature sizes move into deep submicron and into the sub-20 nm range, pattern replication with photolithography becomes increasingly complex and expensive. In recent years, a new pattern replication technique, widely known as nanoimprint lithography (NIL) (Chapter 9), has emerged (Chou *et al.*, 1996). NIL is based on mechanical forming of soft or liquid materials at the nanoscale, as shown in Fig. 10.3. Since its inception, NIL has attracted widespread interest due to its ultra-high resolution (Chou *et al.*, 1997; Hua *et al.*, 2004), low cost, and simple equipment set-up and processing steps. Similar to photolithography, NIL requires a master copy that has to be fabricated by initial pattern generation techniques such as electron-beam lithography or focused-ion beam milling.

Many different versions of nanoimprint techniques have been developed over the last 15 years. UV-curable nanoimprint uses a UV-transparent template (e.g., SiO_2) to mold liquid polymer precursor and then convert the precursor into a cured polymer solid by flood UV exposure (Haisma *et al.*, 1996; Colburn *et al.*, 1999). To improve process throughput, a step-and-flash imprint lithography has been developed (Colburn *et al.*, 1999). Other variations of UV-nanoimprint schemes include combined-nanoimprint-and-photolithography (CNP) (Cheng and Jay Guo, 2004a) and hybrid-mask-mode

(HMM) technique (Cheng and Jay Guo, 2004b). The CNP technique addresses the issues of the residual layers in conventional nanoimprint, and the HMM technique eliminates the difficulty in patterning structures with large size variations by nanoimprint. Due to its low pressure (~0.1 atm) and room temperature processing, UV nanoimprint is more compatible with microelectronic fabrication. Large area patterning of metal interconnect in microelectronic back-end processing has been successfully demonstrated (Chao *et al.*, 2008).

Thermal nanoimprint is mainly used to pattern thermoplastic polymer thin films. It requires heating the polymer thin film to above its glass transition temperature, and a large pressure is applied to emboss the polymer into mold shapes due to the viscous nature of polymer melt. Due to higher pressure, high temperature, and long thermal cycling time, thermal nanoimprint is not suitable for microelectronic fabrication. However, thermal nanoimprint is indispensable in many applications; for example, surface patterning of polymer films or sheets by thermal nanoimprint render unique surface properties such as anti-reflection and superhydrophobicity (Roach *et al.*, 2008). Polymer-based microfluidics and nanofluidic devices can be easily made by thermal nanoimprint for biomedical applications (Truskett and Watts, 2006). And last but not least, thermal nanoimprint is ideal for nondestructive functional polymer patterning, such as direct patterning of conjugated polymers for organic electronics (Cui *et al.*, 2008).

Despite the fast development of nanoimprint technique in the past decade, issues remain for both UV and thermal nanoimprint for industrial adoption. Challenges for commercial manufacturing include throughput improvement, defect mitigation, high-resolution template fabrication, and high-accuracy pattern alignment, to name just a few. Adhesion between the template and the resist layer remains a key issue that limits process throughput and affects defect reduction, particularly for UV nanoimprint. For thermal nanoimprint, both pattern instability caused by residue stress and imprint-induced optical birefringence are additional problems. Ubiquitous chain orientation in polymer nanostructures enforced by nanoimprint is not thermodynamically stable, and the relaxation of the polymer chain can lead to pattern degradation over time. Optical birefringence in polymer nanostructures due to flow-induced chain orientation and the residue stress may also affect certain applications in photonics.

Another useful nanoscale mechanical forming technique is nanoindentation. Using a hard and sharp tip, nanoscale holes and lines can be drawn in both soft and hard materials. Nanoindentation is often used for the characterization of mechanical properties at nanoscale. Particularly, heated AFM tips have been employed to write dot patterns in a flat polymer thin film, such as the famous IBM's 'Millipede' project (Vettiger *et al.*, 2002). The mechanically formed hole pattern in poly(methyl methacrylate) (PMMA) thin film can be used as a platform for high-density information storage. However, similar to thermal nanoimprint, the polymer nanostructures are not stable and lead to long-term stability issues.

10.4 Methods to create nanostructures: bottom-up fabrication of nanostructures

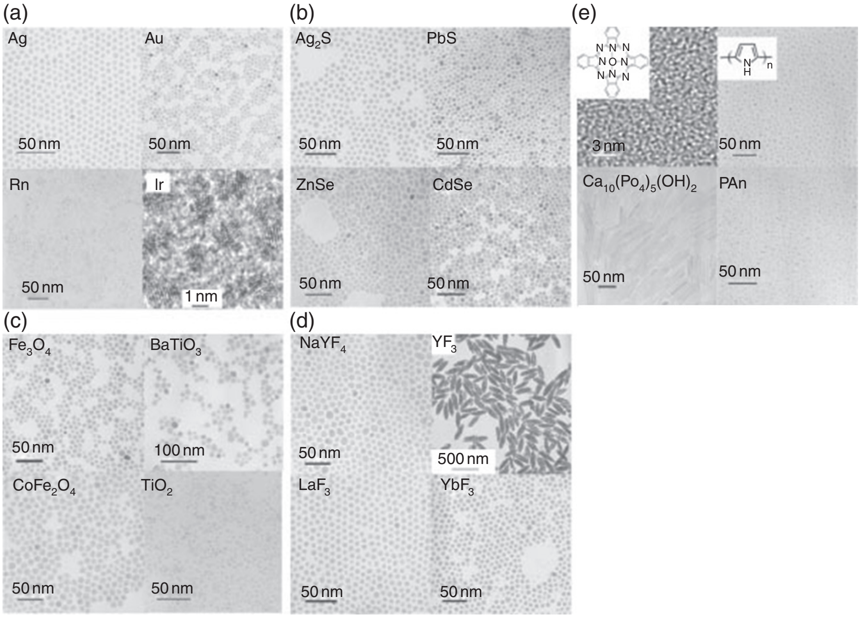
The ‘top-down’ fabrication of nanostructures enjoys the benefits of decades of intensive research and development in microelectronic industry. Those sophisticated tools enable ultra-high resolution (sub-10 nm) patterning with precise pattern placement and registration. However, the extremely high cost of those techniques reserves them for high-end applications only. In many engineering applications, large quantities of nanostructures and nanomaterials are desired. To address this need, many bottom-up techniques have been developed. Those techniques have the capability to produce materials and structures at the nanoscale in large quantities. Well-known bottom-up techniques include solution synthesis, and physical and chemical vapor deposition, among others.

10.4.1 Solution-based synthesis of colloidal nanomaterials

Colloidal science deals with the dispersion of small particles in liquid phase, most commonly water. It has long been an important area of study with diverse technical applications. Nanomaterials dispersed in solution are basically a colloidal system, and the popularity of nanotechnology has reinvigorated research in colloidal science and engineering.

Solution-based synthesis of nanomaterials relies on the precipitation of nanometer-sized crystals from solution after reaction under controlled environment. For example, ZnO nanoparticles can be synthesized by mixing zinc acetate solution with melamine solution and heating to 95°C. However, nanocrystals have large surface areas and are inherently unstable. They can grow into microscale crystals or coagulate into larger aggregates, and eventually precipitate from the solution. Thus, a critical feature in solution-based synthesis is the use of surfactant to limit the growth of the nanomaterials. The surfactant also helps the dispersion of nanomaterials in solvent to form stable suspension.

Solution-base synthesis of nanomaterials is a very versatile technique. Many metallic, semiconducting, and magnetic nanomaterials can be synthesized by this route (Cushing *et al*., 2004; Wang *et al*., 2005), as shown in Fig. 10.4. Also, the processing conditions and the choice of different surfactants enable fine-control of the synthesis. There have been enormous efforts


(a)
Ag
50 nm
Au
50 nm
Rn
50 nm
Ir
1 nm
(b)
Ag_2S
50 nm
PbS
50 nm
ZnSe
50 nm
CdSe
50 nm
(e)
N N N
N O N
N N N
3 nm
N
H
n
50 nm
$Ca_{10}(Po_4)_5(OH)_2$
50 nm
PAn
50 nm
(c)
Fe_3O_4
50 nm
$BaTiO_3$
100 nm
$CoFe_2O_4$
50 nm
TiO_2
50 nm
(d)
$NaYF_4$
50 nm
YF_3
500 nm
LaF_3
50 nm
YbF_3
50 nm

10.4 TEM images of solution-synthesized nanoparticles and nanorods. (a) metallic nanocrystals; (b) compound semiconductor nanocrystals; (c) magnetic and dielectric nanocrystals; (d) rare earth fluorescent nanocrystals; and (e) phosphate and organic nanocrystals. (*Source:* Reprinted with permission from (Wang, Zhuang *et al.* 2005). Copyright: Nature Publishing Group (2005).)

in chemistry and chemical engineering to develop appropriate reaction systems for nanocrystal synthesis. The progress in solution-based production of nanocrystals has enabled the synthesis of nanomaterials of various sizes, shapes, crystal orientations, and compositions (Cushing *et al.*, 2004). The morphologies of the nanomaterials can be uniform, or have a controlled distribution range (Murray *et al.*, 2000). For certain nanomaterials, nanoscale morphologies can be obtained in solution by simply dispersing the bulk materials, such as graphene (Park and Ruoff, 2009) and clay (Möller *et al.*, 2010) nanoplatelets.

The biggest advantage of solution-base synthesis is its low cost. This is particularly important for applications that require a large amount of nanomaterial or large substrate surface areas. Solution-based nanomaterials can be easily applied to large substrate area by techniques such as dip coating or spray coating. Various types of printing techniques, such as inkjet printing, can also be employed for patterning the nanomaterials. It is for those reasons that solution-based colloidal nanomaterials are avidly pursued for applications such as photovoltaic panels and low cost flexible electronics (Talapin *et al.*, 2009).

Core-shell structures are often employed to enable stable and low-defect nanomaterials to improve their electrical and optical properties (Caruso, 2001). Nanomaterials have large surface areas, on which dangling bonds are abundant. The surface of nanomaterial often needs passivation, forming the well-known core–shell structure. The core is typically a semiconducting nanomaterial, while the shell can be another semiconducting material or ligand groups. The shell helps improve the electrical and optical properties of the nanomaterials by lowering the density of surface chemical bonding defects, which often contribute to unwanted defect electronic states within the energy bandgap. Semiconducting nanocrystals with surface passivation have achieved very high fluorescent quantum efficiency. One of the most attractive applications of semiconducting nanoparticles is using them as highly efficient and stable fluorescent tracers in live cells for *in vivo* imaging and diagnostics (Michalet *et al*., 2005). However, the introduction of the shell may impede the transport of electrons and holes among neighboring nanomaterials, resulting in low carrier mobility and hampering the performance of solid-state devices made from those ‘capped’ nanomaterials. This has been one of the major factors that has limited the practical applications of semiconducting nanomaterials in electronic and optoelectronic devices. Also, low temperature processing often leads to crystalline structural defects, and solution processing leads to the incorporation of impurities, both presenting a negative impact on material performance.

10.4.2 Nanomaterials from vapor deposition

Not all nanomaterials can be synthesized from solution processing. Physical and chemical vapor deposition techniques are another group of important techniques to form nanostructured materials (Fang and Zhang, 2006). In physical vapor deposition, source materials to be deposited are first vaporized at high temperature and then condensed on a substrate. During condensation, nanoscale particles, wires, or films can be formed by the vapor–liquid–solid nucleation process. Many elemental and compound semiconductor nanostructures can be easily grown by physical vapor deposition. In chemical vapor deposition, gas precursors are decomposed, or react with other molecules, to produce nanocrystals. The substrate is usually held at elevated temperature to facilitate the chemical reaction and the crystal growth. In many cases, catalyst needs to be deposited onto the substrate first in order to grow the nanomaterials. Chemical vapor deposition is an important technique in the synthesis of the most popular carbon-based nanomaterials, such as fullerene, CNTs, and graphene.

In vapor phase deposition, new morphologies such as nanobelts, nanoribbons and branched nanostructures can also be obtained (Wang, 2004). For nanowires, vapor phase deposition usually ends up with highly directional growth. In most cases, the growth is perpendicular to the substrate surface, although process control such as gas flow direction (Dai and Huang, 2004) or applied electric field (Zhang *et al.*, 2001) can yield lateral growth patterns. The dense vertical nanowire array is an interesting and useful morphology in many device applications. For graphene, the current main objective is to fabricate large area, uniform, and defect-free single-layer graphene sheet on a substrate for device applications (Li *et al.*, 2009).

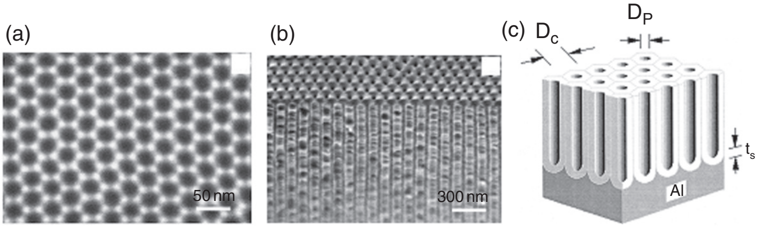
Vapor phase deposition produces nanomaterials with high quality because high temperature is conducive to the formation of high quality crystals, and the gas sources can be easily purified to achieve high purity. The high temperature processing, however, makes vapor deposition incompatible with plastic flexible substrates. Also, the need to control vacuum or gas environment lowers the fabrication throughput.

In addition to the controlled synthesis of nanomaterials, there have been tremendous research activities in arranging those nanomaterials into ordered structures for device applications. In chemical vapor deposition, ordered assembly of nanomaterials can be achieved through careful manipulation of growth parameters. The placement of the nanomaterials on the substrate is usually achieved by lithography-defined template patterns. For nanomaterials dispersed in solution, ordered assembly or super-lattice is usually achieved by placing nanomaterials into prefabricated templates with the help of various forces, such as electric field, magnetic field, and flow-induced orientation. Assembly of nanomaterials using DNA has also been extensively experimented with to achieve higher order structures (Ding *et al.*, 2010).

10.4.3 Miscellaneous bottom-up fabrication techniques

Solution-based synthesis of nanomaterials and vapor deposition of nanomaterials are the two most common techniques in bottom-up fabrication of nanostructures, particularly for semiconducting and metallic nanocrystals. There are a few other very interesting techniques that are capable of forming highly regular nanostructures without elaborate lithographic processes. Two of the most well-known techniques are anodized aluminum oxide (AAO) and block-copolymer (BCP) self-assembly (Chapter 8). Both techniques have been extensively experimented with to achieve highly ordered nanoscale structures, which can be used as sacrificial template towards highly ordered nanostructures in other materials.

When an aluminum thin film is soaked in an acid solution with a suitable substrate bias, its surface will be anodized to form oxide. Due to the electric

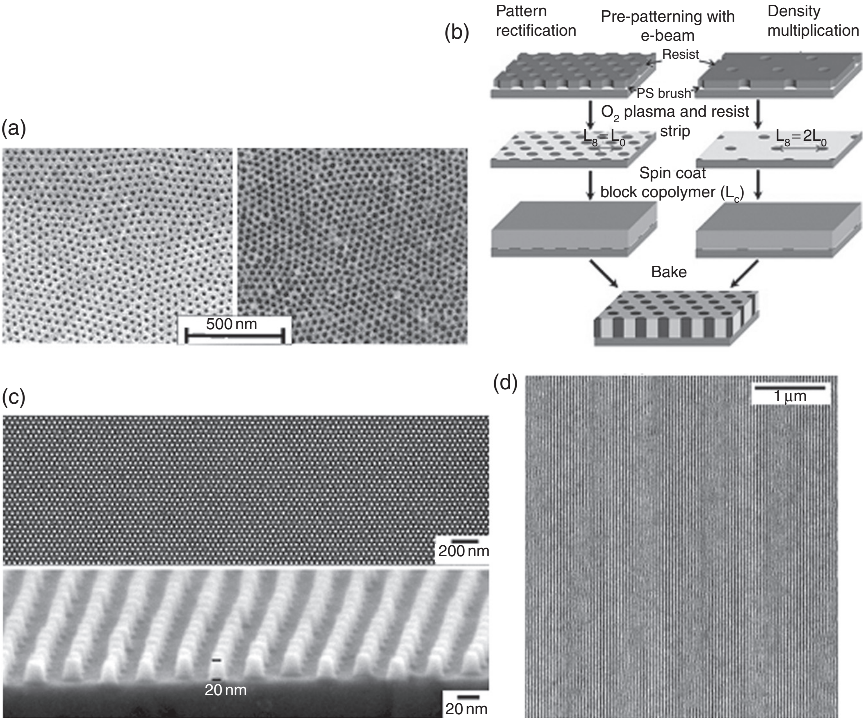


10.5 (a) Top and (b) side view of an AAO template. (*Source:* Reprinted with permission from (Chik and Xu 2004). Copyright: Elsevier (2004).) (c) A schematic of the highly ordered AAO template. (*Source:* Reprinted with permission from (Li, Zhang *et al.* 1998). Copyright: ACS (1998).)

field produced by substrate bias and the competition for aluminum consumption from adjacent oxidation sites, highly ordered holes are typically formed in 2D hexagonal arrangement (Li *et al.*, 1998), as shown in Fig. 10.5. The holes will grow perpendicularly to the surface without widening, due to the self-limited hole growth mechanism, and the hole depth can vary from hundreds of nanometers to hundreds of microns. The diameter of the holes and the periodicity of the hole array can be tuned by acid composition, concentration, solution temperature, and substrate voltage. AAO is unique in its ability to form nanoscale hole arrays with exceptionally large aspect ratios, which could not be achieved otherwise using conventional lithography and dry etching techniques.

There are many applications for AAO thin films. The AAO membranes are typically used as filters to remove particles from solutions. More importantly, AAO membrane can be used as sacrificial template in the fabrication of periodic nanostructures in other functional materials. For example, AAO template has been filled with metals, magnetic alloys, CNT, TiO_2, and ZnO. After removing the AAO template by chemical etching, the magnetic nanodots can be used as high-density bit-patterned media and the CNT, TiO_2 or ZnO nanowires can be used as porous electrodes to enhance the performance of many devices, such as dye-sensitized solar cells, micro fuel cells, and batteries.

Another technique to achieve highly ordered periodic nanostructures is the self-assembly of BCP (Rachel, 2005; S.B, 2007) (Fig. 10.6). BCP is usually synthesized to have two or three incompatible segments that would form phase-segregated structures after thermal or solvent annealing. The annealing provides polymer chain mobility to allow the formation of well-separated phase domains. The size of the domain can be controlled by the length of BCP chain. Unlike the AAO technique, where only a periodic hole pattern can be formed, the final periodic pattern of the anneal BCP thin film varies from periodic dot arrays to well-defined nanoscale gratings.


(a)
500 nm
(b)
Pattern rectification
Pre-patterning with e-beam
Density multiplication
Resist
PS brush
O2 plasma and resist strip
L8 = L0
L8 = 2L0
Spin coat block copolymer (Lc)
Bake
(c)
200 nm
20 nm
20 nm
(d)
1 µm

10.6 (a) An SEM micrograph of disordered dot array obtained by BCP phase segregation and self-assembly (left) and the BCP pattern is transferred into silicon nitride layer by RIE (right). (*Source:* Reprinted with permission from (Park, Harrison *et al*. 1997). Copyright: AAAS (1997).) (b) A schematic of BCP self-assembly on lithographically defined patterns. (c) Highly ordered dot array achieved by the graphoepitaxy technique outlined in (b). (*Source:* Reprinted with permission from (Ruiz, Kang *et al*. 2008) Copyright: AAAS (2008).) (d) Perfectly ordered nanogratings obtained by graphoepitaxy of BCP thin film in lithographically defined 5 μm surface chemical pattern. (*Source:* Reprinted with permission from (Ouk Kim, Solak *et al*. 2003). Copyright: Nature Publishing Group (2003).)

The desired pattern can be tuned by the ratio of the two segments in the BCP main chain, the surface pattern of the substrate and the surface chemistry of the substrate (Rachel, 2005; S.B. 2007). One of the issues for BCP self-assembly is the high density of defects when a large area is needed. To address this issue, extensive study of BCP chain structure, substrate surface pattern, and substrate chemistry (Ouk Kim *et al*., 2003; Bita *et al*., 2008; Ruiz *et al*., 2008) has led to the capability to form sub-20 nm structures over a few centimeters to even a whole-wafer substrate in recent years. This achievement shows the great promise of BCP in future fabrication of high-density and large area periodic nanostructures for integrated circuits and bit-patterned media fabrications. The self-assembly of BCP can be used to fabricate nanoimprint templates, which can then be used to replicate the nanostructures in wafers or substrates in volume production using NIL. Similar to AAO, the high-density BCP nanostructures will be used as template towards the fabrication of metallic, semiconducting, and magnetic nanostructures.

10.5 Properties of nanostructures

Nanostructures have attracted enormous attention in recent years, not only because of their small dimensions for high-density integration but also because of their novel physical properties that are not present in bulk materials. Thus, nanostructures are important for both fundamental research and practical application. The most obvious difference between nanostructures and bulk materials is their surface-to-volume ratio. As the size of a particle reduces, the surface-to-volume ratio increases because this ratio is inversely proportional to the characteristic length of the particle. More atoms reside at the surfaces of the nanostructures than those of bulk materials, so the surface properties become much more pronounced than the bulk properties in nanostructures. For example, nanostructures have a much lower melting point than bulk materials. The large surface-to-volume ratio of nanostructures is attractive in many engineering applications where large surface area is desired, such as surface absorption and catalysis.

Another major difference between nanostructures and bulk materials is the strong confinement of electrons in nanostructures. Such confinement dramatically changes the electronic structure, the transport behavior of the electrons, and the electron–electron interaction. Novel properties arise due to the confinement effects. Detailed analysis of the electrical, optical, thermal, and magnetic properties of the nanostructures can be found in many review articles and textbook chapters (Kittel and McEuen, 2005; Madou, 2011). A few key properties of nanostructures are briefly summarized here.

10.5.1 Electrical and optoelectronic properties of nanostructures

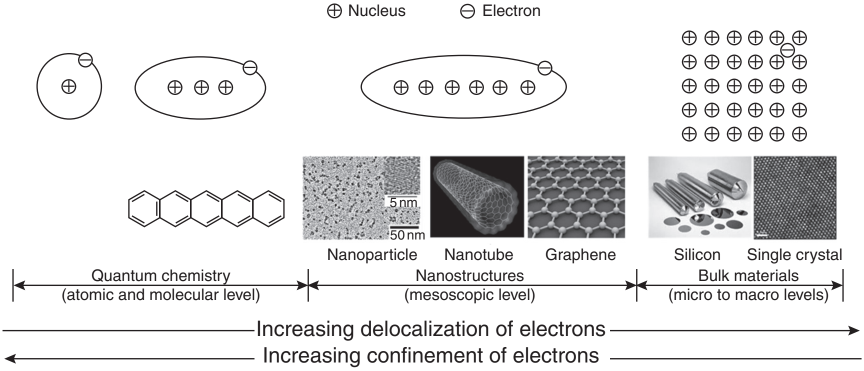
Quantum confinement, electronic structure and density of states

In a crystalline material, the electronic structure is determined by the electronic structure (or atomic orbitals) of the constituent atoms, the crystal symmetry, and the lattice constant. Electrons in conduction band are delocalized

Table 10.1 Comparison of nanostructures and bulk materials

Material	Dimensions of confinement	Dimensions of material	Density of states
Nanoparticle	3	0D	2
Nanowire	2	1D	$\frac{(2\pi m*)^{1/2}}{h}E^{-1/2}$
Nanowell	1	2D	$\frac{4\pi}{h^2}m*$
Bulk	0	3D	$\frac{8\sqrt{2}\pi}{h^3}(m*)^{3/2}E^{1/2}$

E is the energy above the band-edge or ground energy level; *m** is the effective mass of the electrons.


⊕ Nucleus
⊖ Electron
5 nm
50 nm
Nanoparticle
Nanotube
Graphene
Silicon
Single crystal
Quantum chemistry (atomic and molecular level)
Nanostructures (mesoscopic level)
Bulk materials (micro to macro levels)
Increasing delocalization of electrons
Increasing confinement of electrons

10.7 A schematic of electron confinement or delocalization from atom to bulk materials.

in infinite periodic potentials. In nanostructures, one or more dimensions have limited sizes, forming confinement for electrons along those directions (Table 10.1). In nanowell structures, electrons are confined in one dimension and free in the other two dimensions. Such a structure is called a two-dimensional (2D) nanostructure. In one-dimensional (1D) nanostructures, such as nanowires, electrons are confined in two dimensions but free in one dimension. In zero-dimensional (0D) nanoparticles, electrons are confined in all three dimensions. At the extreme end, further confinement of electrons in even smaller dimensions enters the realms of molecules and atoms. The trend in electron confinement or delocalization from atoms and molecules to nanostructures and bulk materials is schematically shown in Fig. 10.7.

In bulk materials, there are a large number of allowed electron wavefunctions (also called energy states or crystal orbitals) in the valence and conduction bands. The energy distribution of those energy states is continuous within those bands. The electron confinement in nanomaterials, however, results in a very limited number of available energy states for electrons to occupy. More importantly, the corresponding energies of those states in nanomaterials are no longer continuous as those in bulk materials. The electronic structures of those reduced-dimension nanostructures can be solved using the simple 1D, 2D, and 3D 'particle-in-a-box' models for quantum wells, nanowires, and quantum dots, respectively. The density of states for those nanostructures has very different energy dependency from bulk materials, due to different level of confinement, as shown in Table 10.1.

In 0D nanoparticles, the electronic states have discrete energy levels, and each can accommodate a finite number of electrons, determined by the Pauli's exclusion principle and the degeneracy of the energy states. The density of states of 0D nanoparticles, therefore, are sharp peaks centered at the allowed energy levels. Thus, the electronic structure of a quantum dot is very similar to an atom or a molecule. In this sense, a quantum dot can be regarded as either an artificial 'atom' or a large 'molecule.' Advanced quantum dot nanostructures, such as core-shell structures, can be fabricated to further manipulate energy states in those artificially engineering nanomaterials.

Electronic transport in nanostructures

Due to their unique electronic structures and density of states, nanostructures have new properties that are not present in bulk materials. Novel electronic and optoelectronic devices can be fabricated to take advantage of the unique properties of the nanostructures.

In all nanostructures, 2D quantum well structure is the easiest to fabricate. This is because there is only one confined dimension at the nanoscale and it can be the thickness of a thin film, which can be precisely controlled with mature thin film deposition techniques such as molecular beam epitaxy or chemical vapor deposition. By growing nanometer-thick narrower bandgap semiconductors between wider bandgap semiconductor layers, a quantum well structure can be easily formed. The confinement in the film growth direction leads to the appearance of discrete energy sub-bands in the conduction and valence bands of the narrower bandgap semiconductors.

Electron transport through stacked quantum well structures is dictated by the tunneling process. The alignment of the electron energy levels between neighboring quantum wells dramatically changes the tunneling probability. By applying a voltage to tilt the electronic bands in serial quantum wells, resonant tunneling can occur at several voltage levels. The current-voltage characteristic of such resonant tunneling diode has a unique negative differential resistance region, which is useful for high-frequency microwave devices.

In 1D nanowire, electron transport also has a new feature – the quantization of conductance. The total current flowing through a nanowire biased at voltage V is given as: $I = 2e^2V/h$, where e is electron charge and h is the Planck's constant. This I–V characteristic is a consequence of inverse dependence of density of states on electron velocity. The term $2e^2/h$ is called the conductance quantum. The inverse of the conductance quantum is called the resistance quantum, and its value can be calculated as 12.906 kΩ. This result indicates that, even without any scattering events, a 1D nanowire always has finite conductance. The quantization of conductance has been well observed in experiments (van Wees *et al.*, 1988).

In 0D quantum dots, the atom-like electronic structure presents a unique feature. Electrons can be individually added to the quantum dots, in a way similar to the electron filling of atomic orbitals in an atom. For the extra electron to be added into the nanoparticle, it must overcome the required charging energy. This charging energy is determined by both the electron–electron repulsion and the next available energy state for the extra electron to occupy. When a nanoparticle is sandwiched between two contacts, electrons can tunnel through the nanoparticle if the energy state of the nanoparticle is lowered through a gate contact. Otherwise electrons will not be able to tunnel through the nanoparticle, due to the Coulomb blockade effect. Such a device is called a single-electron transistor (SET) and it is very attractive because the current comes from the movement of one electron at a time. For SET to operate at room temperature, the electron charging energy must be greater than the thermal energy. This requires extremely small nanoparticles, typically in the range of a few nanometers. Despite extremely challenging fabrication, SET still attracts a lot of interest in the microelectronic community, because it enables the processing of electronic signal using individual electrons instead of an assembly of a large number of electrons in current MOSFET devices. In principle, extremely power-efficient SETs can be constructed for future high-density and low-power integrated circuits.

Optoelectronic properties of nanostructures

Many optical processes, such as light absorption and emission, in materials are directly related to the electronic structure of the material. Since nanostructures have unique electronic structure and density of states, their optical properties are also dramatically different from bulk materials, enabling a range of novel devices or improved devices for optoelectronic applications.

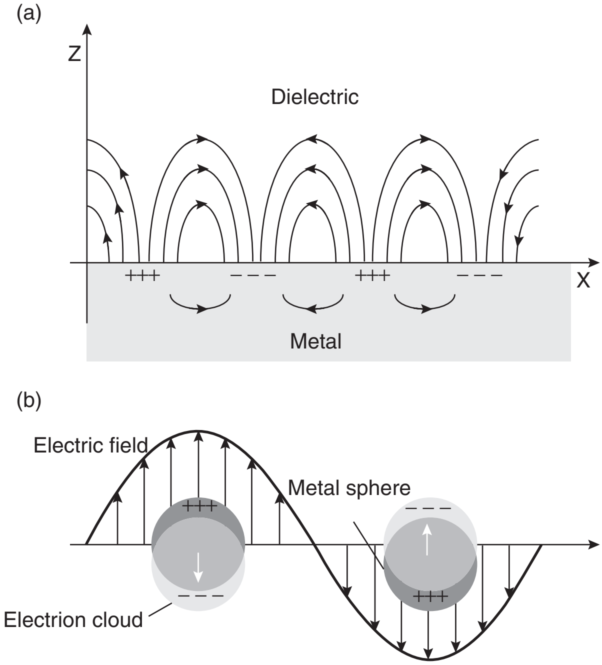
In 2D quantum well structures, the appearance of discrete energy levels along the confinement direction enables new intra-band and inter-band electronic transitions, leading to modified optical absorption and emission spectra. A very attractive application is for the emission of longer infrared wavelengths (intra-band transition) and shorter wavelengths (inter-band transitions other than band-edge transitions). Quantum well structures are also very useful in infrared detection by taking advantage of the intra-band transitions. When multiple quantum wells are stacked together, quantum cascade light emission that results in high-power mid-infrared lasing is achieved (Faist *et al.*, 1994; Claire *et al.*, 2001). Currently, 2D quantum well structures are widely used in optoelectronic devices such as LEDs, solid-state lasers, and photodetectors (Levine, 1993; Morkoc *et al.*, 1994).

Semiconducting 1D nanostructures, such as ZnO nanowires prepared by vapor phase epitaxial growth, can function as an ultra-compact waveguide. The two ends of the nanowire are flat crystalline facets, which can act as two dielectric mirrors to form a built-in Fabry–Pérot resonator. This feature allows the single nanowire to function as an ultra-compact laser (Huang *et al.*, 2001), which may find many applications in lithography, data storage, and optical information processing.

In 0D quantum dots, the difficulty for carriers, electrons, and holes to hop onto neighboring dots provides strong localization of electron-and-hole pairs within the quantum dot. Because of the small size of the nanoparticle, the exciton formed by the electron-and-hole pair has very large binding energy, and thus a very high recombination efficiency. This feature renders quantum dots efficient fluorescent materials with very high quantum efficiency. In solid-state devices, highly efficient light emission leads to quantum dot lasers with low threshold current (Huffaker *et al.*, 1998). In solution-processed quantum dots, those quantum dots can be used as fluorescent tags for bio-applications and light emitters in solution-processed electroluminescent devices.

10.5.2 Optical properties of metallic nanostructures

Metallic nanostructures have the capability to strongly couple incident electromagnetic radiation into SP modes (Ritchie, 1957; Raether, 1988). SPs are collective electron oscillations at certain metal–dielectric interfaces. Surface plasmon polaritons (SPP) can be formed when SPs are excited by incident electromagnetic waves and coupled with photons. The two most significant features of the SP modes are (1) exceptionally large field enhancement, and (2) extremely tight spatial confinement of light energy. These distinctive characteristics of SPs have spurred the research in plasmonics for a broad range of applications (Gramotnev and Bozhevolnyi, 2010; Schuller *et al.*, 2010). Plasmonic resonance in metallic nanostructures at the visible and near-IR wavelengths can lead to dramatic field enhancement due to their small structures. The amplified light–matter interaction is beneficial to applications where strong light intensity is desired, such as chemical and bio-detection with optical techniques for trace level chemicals, photodetectors,



10.8 Schematics of surface plasmon polaritons excited by light: (a) propagating surface plasmon at a smooth metal–dielectric interface, and (b) localized surface plasmon on metal nanoparticle surface. The + and – signs represent low and high electron densities. (*Source:* Reprinted with permission from (Willets and Van Duyne 2007). Copyright: Annual Reviews (2007).)

and solar cells. The extremely small mode volume of the SPs is ideal for scanning-based ultra-high-resolution imaging and lithography to overcome the resolution limitation introduced by diffraction in conventional optical components.

Depending on their spatial distribution, there are localized surface plasmons (LSP) and propagating surface plasmons (PSP), as shown in Fig. 10.8. LSPs usually exist on isolated metallic nanostructures that are much smaller than the wavelength of the excitation. They are strongly confined electron oscillations that cannot propagate. The resonant wavelength of the LSPs strongly depends on the shape, size, and dielectric functions of both the metal and the dielectric surroundings (Chen *et al*., 2007; Jain *et al*., 2008). Active research in this area involves the design and fabrication of various LSP structures at different resonant wavelengths and with maximum field enhancement. The strongly confined optical energy at the sharp tip, edge, or corner of the nanostructures can be used for nonlinear spectroscopy and nanolithography. The most prominent applications are substrates for surface-enhanced Raman spectroscopy and scanning probe tips for optical lithography at the deep subwavelength scale.

PSPs usually reside on continuous or coupled metallic nanostructures. The propagation of PSPs along the metal–dielectric interface guide the transport of optical energies; thus an attractive application of the PSPs is the highly compact optical waveguide. However, the propagation of the PSPs is accompanied by significant energy loss, resulting in limited propagation lengths. Novel design is still desired to minimize propagation loss for practical waveguide applications.

Plasmonic structures greatly reduce the dimensions of photonic integrated circuits through their ability to manipulate electromagnetic waves at the nanoscale. They have the potential to realize optical interconnects for electronic devices and ultra-compact plasmonic chips for information processing (Ozbay, 2006). The combination of plasmonic nanostructures with conventional electronic devices presents new paradigms of monolithic integration of photonic and electronic components. For these reasons, many exciting novel applications of plasmonics are expected in the near future.

10.6 Applications of nanostructures

The unique features in nanostructures, such as large surface area, energy level discretization due to quantum confinement, novel density of states and phonon coupling, and capability of exciting surface plasmonic modes, have given nanostructures distinctive properties that are usually absent in their bulk counterparts. More importantly, many of those unique properties can be fine-tuned through structural design, providing flexibility in tailoring specific properties for target applications. The potential application areas for nanostructures are extremely broad, from the microelectronics industry to bioengineering, because nanostructures are truly a multidisciplinary topic.

In each section, we have discussed specific applications for each nanostructure and for each unique property. Some of those applications are already seen in commercial products, while others remain in the lab. Due to the maturity of top-down fabrication techniques, lithographically defined nanostructures have been widely used in nanoelectronics for achieving better device, circuit, and system performance, as demonstrated in modern day CPUs and memory devices. For optoelectronic devices, quantum dots and quantum wells grown by MBE or MOCVD techniques have led to a series of breakthroughs in solid-state lasers, photodetectors, and high-frequency resonant tunneling devices. Other areas, such as high-density magnetic information storage, are being hotly pursued, with the main focus on the perfection of fabricating large area and high-density bit-patterned media with extremely low defect density and high fabrication throughput.

Photonic nanostructures also show great potential in practical applications. Surface-enhanced Raman scattering (SERS) based on metallic nanostructures are becoming a standard technique for detecting trace level molecules. The strong field enhancement and tight spatial confinement of optical energy are promising for nonlinear optics, high-resolution imaging, deep subwavelength optical lithography, and energy focusing. In additional to fundamental research, current efforts in these areas include developing an easy-to-use plasmonic platform for commercial applications.

Nanoparticles and nanowires synthesized in solution or deposited by vapor deposition have attracted great interest for their easy processing, low cost, and high throughput. Those carbon-based and compound semiconductor nanomaterials are currently being explored for energy harvesting, biomedical engineering, catalysts, and new nanocomposite materials. However, other issues, such as the limitations informing higher order nanostructures, have seriously limited their real potential. Particularly for electronic applications, there exist several prominent challenges: the need for the alignment and placement of the nanomaterials for integrated circuit fabrication, the practical performance deterioration introduced by large density of defects, surface states and impurities, and the interference of charge transport by the often-needed surface passivation groups. There have been high hopes for those semiconducting nanostructures, but then the practical issues outlined above kill their chances, such as in the case of CNTs. Although there is a strong resurgence in another carbon-based nanomaterial – graphene – its real potential in the microelectronic industry may be very limited due to many practical issues.

It is interesting to investigate the new science in nanostructures, but the limitations of nanostructures in engineering applications should not be overlooked. Often it is not enough to just show that nanostructures can be made into certain devices or can achieve certain functionalities; we also need to compare holistically between nanostructures and conventional techniques in terms of functionality and cost. For engineering applications, a superior material property is often not enough to deliver its success in commercial applications. Also, in recent years, the health and environmental impact of some nanomaterials have raised concerns. Thus, a candid and objective evaluation of the value of some nanostructures in practical applications should be followed, to make the best out of research efforts and resources.

10.7 References

Avouris, P., Chen, Z.and Perebeinos, V. (2007). 'Carbon-based electronics'. *Nature Nanotechnology*, **2**(10): 605–615.

Bita, I., Yang, J. K. W., Jung, Y. S., Ross, C. A., Thomas, E. L. and Berggren, K. K. (2008). 'Graphoepitaxy of self-assembled block copolymers on two-dimensional periodic patterned templates'. *Science*, **321**(5891): 939–943.

Brueck, S. R. J. (2005). 'Optical and interferometric lithography – nanotechnology enablers'. *Proceedings of the IEEE*, **93**(10): 1704–1721.

Caruso, F. (2001). 'Nanoengineering of particle surfaces'. *Advanced Materials*, **13**(1): 11–22.

Chao, B. H., Palmieri, F., Jen, W.-L., McMichael, D. H., Willson, C. G., Owens, J., Berger, R., Sotoodeh, K., Wilks, B., Pham, J., Carpio, R., LaBelle, E. and Wetzel, J. (2008). 'Dual damascene BEOL processing using multilevel step and flash imprint lithography'. *Proceedings of SPIE – The International Society for Optical Engineering*, **6921**(1): 69210C.

Chen, J., Wiley, B. J. and Xia, Y. (2007). 'One-dimensional nanostructures of metals: Large-scale synthesis and some potential applications'. *Langmuir*, **23**(8): 4120–4129.

Cheng, X. and Jay Guo, L. (2004a). 'A combined-nanoimprint-and-photolithography patterning technique'. *Microelectronic Engineering*, **71**(3–4): 277–282.

Cheng, X. and Jay Guo, L. (2004b). 'One-step lithography for various size patterns with a hybrid mask-mold'. *Microelectronic Engineering*, **71**(3–4): 288–293.

Chik, H. and Xu, J. M. (2004). 'Nanometric superlattices: non-lithographic fabrication, materials, and prospects'. *Materials Science and Engineering: R: Reports*, **43**(4): 103–138.

Chou, S. Y., Krauss, P. R. and Renstrom, P. J. (1996). 'Imprint lithography with 25-nanometer resolution'. *Science*, **272**(5258): 85–87.

Chou, S. Y., Krauss, P. R., Zhang, W., Guo, L. and Zhuang, L. (1997). 'Sub-10 nm imprint lithography and applications'. *Journal of Vacuum Science and Technology B*, **15**: 2897–2904.

Claire, G., Federico, C., Deborah, L. S. and Alfred, Y. C. (2001). 'Recent progress in quantum cascade lasers and applications'. *Reports on Progress in Physics*, **64**(11): 1533.

Colburn, M., Johnson, S. C., Stewart, M. D., Damle, S., Bailey, T. C., Choi, B., Wedlake, M., Michaelson, T. B., Sreenivasan, S. V., Ekerdt, J. G. and Willson, C. G. (1999). 'Step and flash imprint lithography: a new approach to high-resolution patterning'. *Proceedings of SPIE – The International Society for Optical Engineering*, **3676**: 379–389.

Cui, D., Li, H., Park, H. and Cheng, X. (2008). 'Improving organic thin film transistor performance by nanoimprint-induced chain ordering'. *Journal of Vacuum Science and Technology B*, **26**: 2404–2409.

Cushing, B. L., Kolesnichenko, V. L. and O'Connor, C. J. (2004). 'Recent advances in the liquid-phase syntheses of inorganic nanoparticles'. *Chemical Reviews*, **104**(9): 3893–3946.

Dai, L. and Huang, S. (2004). Multilayer carbon nanotube films and method of making the same, US Patent No: 6808746.

Darling, S.B. (2007). 'Directing the self-assembly of block copolymers'. *Progress in Polymer Science*, **32**(10): 1152–1204.

Ding, B., Deng, Z., Yan, H., Cabrini, S., Zuckermann, R. N. and Bokor, J. (2010). 'Gold nanoparticle self-similar chain structure organized by DNA origami'. *Journal of the American Chemical Society*, **132**(10): 3248–3249.

Faist, J., Capasso, F., Sivco, D. L., Sirtori, C., Hutchinson, A. L. and Cho, A. Y. (1994). 'Quantum cascade laser'. *Science*, **264**(5158): 553–556.

Fang, X. and Zhang, L. (2006). 'Controlled growth of one-dimensional oxide nanomaterials'. *Journal of Materials Science and Technology*, **22**(1): 1–18.

Ginger, D. S., Zhang, H.and Mirkin, C. A. (2004). 'The evolution of dip-pen nanolithography'. *Angewandte Chemie International Edition*, **43**(1): 30–45.

Gramotnev, D. K. and Bozhevolnyi, S. I. (2010). 'Plasmonics beyond the diffraction limit'. *Nature Photonics*, **4**(2): 83–91.

Haisma, J., Verheijen, M., van den Heuvel, K. and van den Berg, J. (1996). 'Mold-assisted nanolithography: A process for reliable pattern replication'. *Journal of Vacuum Science & Technology B*, **14**: 4124–4128.

Hartschuh, A. (2008). 'Tip-enhanced near-field optical microscopy'. *Angewandte Chemie – International Edition*, **47**(43): 8178–8191.

Hong, S. and Mirkin, C. A. (2000). 'A nanoplotter with both parallel and serial writing capabilities'. *Science*, **288**(5472): 1808–1811.

Hua, F., Sun, Y., Gaur, A., Meitl, M. A., Bilhaut, L., Rotkina, L., Wang, J., Geil, P., Shim, M., Rogers, J. A. and Shim, A. (2004). 'Polymer imprint lithography with molecular-scale resolution'. *Nano Letters*, **4**(12): 2467–2471.

Huang, M. H., Mao, S., Feick, H., Yan, H., Wu, Y., Kind, H., Weber, E., Russo, R. and Yang, P. (2001). 'Room-temperature ultraviolet nanowire nanolasers'. *Science*, **292**(5523): 1897–1899.

Huffaker, D. L., Park, G., Zou, Z., Shchekin, O. B. and Deppe, D. G. (1998). '1.3 μm room-temperature GaAs-based quantum-dot laser'. *Applied Physics Letters*, **73**(18): 2564–2566.

Huo, F., Zheng, Z., Zheng, G., Giam, L. R., Zhang, H. and Mirkin, C. A. (2008). 'Polymer pen lithography'. *Science*, **321**(5896): 1658–1660.

Jain, P. K., Huang, X., El-Sayed, I. H. and El-Sayed, M. A. (2008). 'Noble metals on the nanoscale: Optical and photothermal properties and some applications in imaging, sensing, biology, and medicine'. *Accounts of Chemical Research*, **41**(12): 1578–1586.

Kittel, C. and McEuen, P. (2005). *Nanostructures*. *Introduction to Solid State Physics*. Hoboken, NJ, Wiley: 515–564.

Levine, B. F. (1993). 'Quantum-well infrared photodetectors'. *Journal of Applied Physics*, **74**(8): R1–R81.

Li, F., Zhang, L. and Metzger, R. M. (1998). 'On the growth of highly ordered pores in anodized aluminum oxide'. *Chemistry of Materials*, **10**(9): 2470–2480.

Li, X., Cai, W., An, J., Kim, S., Nah, J., Yang, D., Piner, R., Velamakanni, A., Jung, I., Tutuc, E., Banerjee, S. K., Colombo, L. and Ruoff, R. S. (2009). 'Large-area synthesis of high-quality and uniform graphene films on copper foils'. *Science*, **324**(5932): 1312–1314.

Madou, M. J. (2011). *Solid-State Physics, Fluidics, and Analytical Techniques in Micro- and Nanotechnology*. Boca Raton, FL, Taylor and Francis.

Michalet, X., Pinaud, F. F., Bentolila, L. A., Tsay, J. M., Doose, S., Li, J. J., Sundaresan, G., Wu, A. M., Gambhir, S. S. and Weiss, S. (2005). 'Quantum dots for live cells, in vivo imaging, and diagnostics'. *Science*, **307**(5709): 538–544.

Möller, M. W., Handge, U. A., Kunz, D. A., Lunkenbein, T., Altstädt, V. and Breu, J. (2010). 'Tailoring shear-stiff, mica-like nanoplatelets'. *ACS Nano*, **4**(2): 717–724.

Morkoc, H., Strite, S., Gao, G. B., Lin, M. E., Sverdlov, B. and Burns, M. (1994). 'Large-band-gap SiC, III-V nitride, and II-VI ZnSe-based semiconductor device technologies'. *Journal of Applied Physics*, **76**(3): 1363–1398.

Murray, C. B., Kagan, C. R. and Bawendi, M. G. (2000). 'Synthesis and characterization of monodisperse nanocrystals and close-packed nanocrystal assemblies'. *Annual Review of Materials Science*, **30**: 545–610.

Ouk Kim, S., Solak, H. H. Stoykovich, M. P., Ferrier, N. J., de Pablo, J. J. and Nealey, P. F. (2003). 'Epitaxial self-assembly of block copolymers on lithographically defined nanopatterned substrates'. *Nature*, **424**(6947): 411–414.

Ozbay, E. (2006). 'Plasmonics: Merging photonics and electronics at nanoscale dimensions'. *Science*, **311**(5758): 189–193.

Park, M., Harrison, C. Chaikin, P. M., Register, R. A. and Adamson, D. H. (1997). 'Block copolymer lithography: Periodic arrays of ~10^{11} holes in 1 square centimeter'. *Science*, **276**(5317): 1401–1404.

Park, S. and Ruoff, R. S. (2009). 'Chemical methods for the production of graphenes'. *Nature Nanotechnology*, **4**(4): 217–224.

Piner, R. D., Zhu, J., Xu, F., Hong, S. and Mirkin, C. A. (1999). 'Dip-pen' nanolithography'. *Science*, **283**(5402): 661–663.

Rachel, A. S. (2005). 'Patterning with block copolymer thin films'. *Materials Science and Engineering: R: Reports*, **48**(6): 191–226.

Raether, H. (1988). Surface plasmons on smooth and rough surfaces and on gratings. Berlin; New York, Springer-Verlag.

Ritchie, R. H. (1957). 'Plasma losses by fast electrons in thin films'. *Physical Review*, **106**(5): 874–881.

Roach, P., Shirtcliffe, N. J. and Newton, M. I. (2008). 'Progess in superhydrophobic surface development'. *Soft Matter*, **4**(2): 224–240.

Ruiz, R., Kang, H., Detcheverry, F. A., Dobisz, E., Kercher, D. S., Albrecht, T. R., de Pablo, J. J. and Nealey, P. F. (2008). 'Density multiplication and improved lithography by directed block copolymer assembly'. *Science*, **321**(5891): 936–939.

Salaita, K., Wang, Y. and Mirkin, C. A. (2007). 'Applications of dip-pen nanolithography'. *Nature Nanotechnology*, **2**(3): 145–155.

Schattenberg, M. L., Anderson, E. H. and Smith, H. I. (1990). 'X-ray/VUV transmission gratings for astrophysical and laboratory applications'. *Physica Scripta*, **41**: 13–20.

Schuller, J. A., Barnard, E. S., Cai, W., Jun, Y. C., White, J. S. and Brongersma, M. L. (2010). 'Plasmonics for extreme light concentration and manipulation'. *Nature Materials*, **9**(3): 193–204.

Shim, W., Braunschweig, A. B., Liao, X., Chai, J., Lim, J. K., Zheng, G. and Mirkin, C. A. (2011). 'Hard-tip, soft-spring lithography'. *Nature*, **469**(7331): 516–520.

Sun, S., Chong, K. S. L. and Leggett, G. J. (2002). 'Nanoscale molecular patterns fabricated by using scanning near-field optical lithography'. *Journal of the American Chemical Society*, **124**(11): 2414–2415.

Talapin, D. V., Lee, J.-S., Kovalenko, M. V. and Shevchenko, E. V. (2009). 'Prospects of colloidal nanocrystals for electronic and optoelectronic applications'. *Chemical Reviews*, **110**(1): 389–458.

Truskett, V. N. and Watts, M. P. C. (2006). 'Trends in imprint lithography for biological applications'. *Trends in Biotechnology*, **24**(7): 312–317.

van Wees, B. J., van Houten, H., Beenakker, C. W. J., Williamson, J. G., Kouwenhoven, L. P., van der Marel, D. and Foxon, C. T. (1988). 'Quantized conductance of point contacts in a two-dimensional electron gas'. *Physical Review Letters*, **60**(9): 848–850.

Vettiger, P., Cross, G., Despont, M., Drechsler, U., Durig, U., Gotsmann, B., Haberle, W., Lantz, M. A., Rothuizen, H. E., Stutz, R. and Binnig, G. K. (2002). 'The 'millipede' – nanotechnology entering data storage'. *Nanotechnology, IEEE Transactions on*, **1**(1): 39–55.

Wang, X., Zhuang, J.,Peng, Q. and Li, Y. (2005). 'A general strategy for nanocrystal synthesis'. *Nature*, **437**(7055): 121–124.

Wang, Z. L. (2004). 'Zinc oxide nanostructures: Growth, properties and applications'. *Journal of Physics: Condensed Matter*, **16**(25): R829–R858.

Willets, K. A. and Van Duyne, R. P. (2007). 'Localized surface plasmon resonance spectroscopy and sensing'. *Annual Review of Physical Chemistry*, **58**(1): 267–297.

Willson, C. G. and Roman, B. J. (2008). 'The future of lithography: SEMATECH litho forum 2008'. *ACS Nano*, **2**(7): 1323–1328.

Zhang, Y., Chang, A., Cao, J., Wang, Q., Kim, W., Li, Y., Morris, N., Yenilmez, E., Kong, J. and Dai, H. (2001). 'Electric-field-directed growth of aligned single-walled carbon nanotubes'. *Applied Physics Letters*, **79**(19): 3155–3157.

11

Nanophotonics: devices for manipulating light at the nanoscale

P. DASTMALCHI, A. HADDADPOUR and
G. VERONIS, Louisiana State University, USA

DOI: 10.1533/9780857098757.376

Abstract: This chapter reviews nanophotonic devices for manipulating light at the nanoscale. The chapter first discusses wavelength-scale nanophotonic devices, such as photonic crystals and ring resonators. It then discusses several topics associated with manipulating light at subwavelength scales with nanometallic structures: extraordinary optical transmission through subwavelength apertures, optical nanoantennas, plasmonic focusing, near-field optical microscopy, and plasmonic waveguides. Finally, the chapter discusses the application of nanophotonics in enhancement of nonlinear processes and photovoltaics.

Key words: plasmonics, nanophotonics, photonic crystals, ring resonators.

11.1 Introduction

Nanophotonics is the science and engineering of light–matter interactions at the nanoscale. Dielectric nanophotonic structures and devices, such as photonic crystal devices, enable wavelength-scale manipulation of light. In addition, light-guiding structures that allow subwavelength confinement of the optical mode are important for achieving compact integrated photonic devices (Maier *et al.*, 2003). However, the minimum confinement of a guided optical mode in dielectric waveguides is set by the diffraction limit, and is of the order of λ_0/n, where λ_0 is the wavelength in free space and n is the refractive index. As opposed to dielectric devices, plasmonic devices have shown the potential to guide subwavelength optical modes, the so-called surface plasmon polaritons, at metal–dielectric interfaces.

In this chapter we first discuss wavelength-scale nanophotonic devices, such as photonic crystals (PhC) and ring resonators. We then discuss several topics associated with manipulating light at subwavelength scales with nanometallic structures: extraordinary optical transmission through subwavelength apertures, optical nanoantennas, plasmonic focusing, near-field optical microscopy, and plasmonic waveguides. Finally, we discuss the application of nanophotonics in enhancement of nonlinear processes and photovoltaics.

11.2 Photonic crystals

PhCs are periodic dielectric and metallic optical materials and structures designed to affect the propagation of electromagnetic (EM) waves. They have many analogies to solid-state crystals. Similar to the familiar crystals of atoms, PhCs have discrete translational symmetry. They have spatial periodicity in their dielectric constant.

The most interesting property of PhCs is that they can support photonic band gaps (PBG), which are a range of frequencies where no propagating solutions of Maxwell's equations exist. In other words, PBG in PhCs is a range of forbidden frequencies or wavelengths in which photons with frequency or wavelength lying in the PBG cannot propagate through the medium. Tremendous interest has been drawn to PhCs from both fundamental and practical viewpoints. Photonic crystal structures can be engineered accurately in order to meet the requirements of specific applications.

One-dimensional (1-D) PhCs consist of periodic multilayer structures. Their properties have been extensively investigated theoretically. Fabricating these structures was initially challenging. While 1-D PhCs are useful in many applications, the most interesting properties of PhCs arise once two-dimensional (2-D) and three-dimensional (3-D) structures are considered. As an example, 2-D photonic crystal slabs, which consist of a 2-D periodic structure formed in a dielectric slab, have the capability of controlling photons much more strongly than 1-D crystals. In general, 2-D PhCs consist of periodic arrays with square or hexagonal lattice. These two lattice types usually lead to larger band gap in the band structure in comparison with other lattice types. A 2-D PhC structure consisting of a square lattice of dielectric rods with lattice constant of a and cylinder radius of $0.1a$ is illustrated in Fig. 11.1a. Fabrication of 3-D PhCs is quite challenging, so most PhC applications involve 2-D photonic crystal slabs. Photonic crystal structures possess a variety of band dispersions and band gaps.

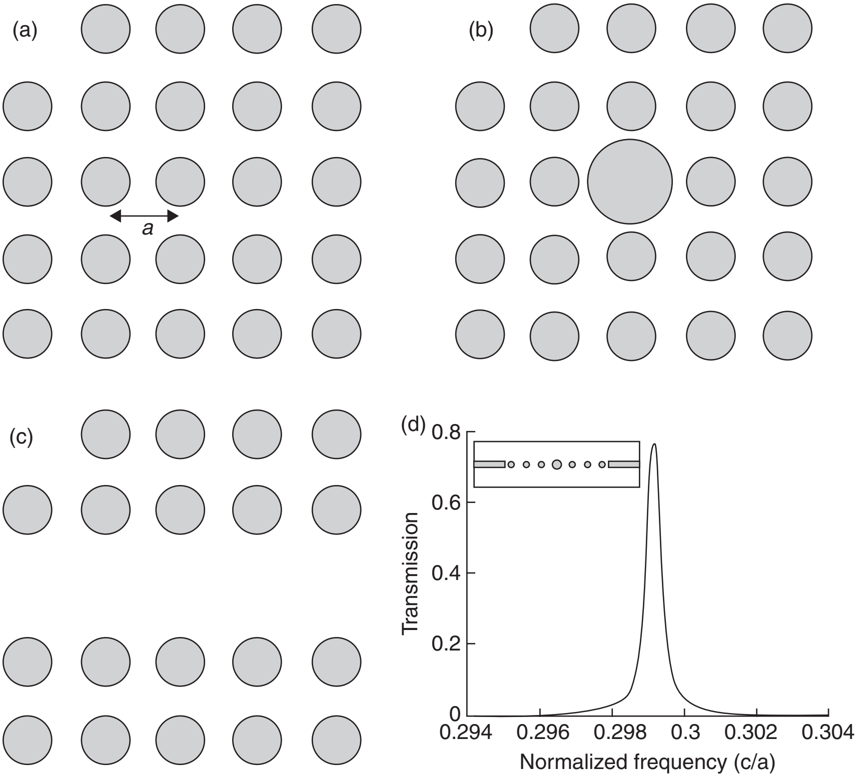
11.2.1 Photonic crystal band gap

The photonic band structure provides us information about the propagation properties of EM radiation within the photonic crystal. Solving Maxwell's equations analytically for two- or three-dimensional periodic lattices is not possible. However, numerical computational techniques, such as the finite-difference time-domain (FDTD) method, finite-difference frequency-domain (FDFD) method, finite integration technique (FIT), plane wave method (PWM), and others can be used to calculate the dispersion of PhCs. The geometry of the chosen structure in 2-D PhCs plays an important role in obtaining PBG for transverse magnetic (TM) and transverse electric (TE) modes. The most challenging issue associated with 2-D PhCs is to design them to have large band gaps. When the PBGs for TM and TE modes overlap, or in other words the TM and TE PBGs are in the same spectral range, they form combined PBGs known as complete band gaps. The existence of a large complete PBG paves the way for designing waveguides, which can support propagating modes of both polarizations in the same frequency spectral range. There have been various attempts and suggestions for enlarging PBGs in PhCs. In a recent work, Proietti Zaccaria *et al*. (2012) showed that metallic PhCs can provide complete PBGs, which is not possible with dielectric PhCs. This characteristic is related to the dispersive properties of the metallic parts of the crystal.

11.2.2 Defects and cavities in photonic crystals

Introducing disorders in the regular structure of PhCs, known as defects, creates highly localized EM modes in the band gap. A defect may result in allowed states for particular photon frequencies in the PBG. Point and line defects are used to achieve ultra-high quality microcavities and waveguides, respectively. Point defects in photonic crystal structures can be obtained by removing a rod (hole defect), and altering the radius or refractive index of the rod in comparison with the rest of the structure. Similarly, simply applying the above-mentioned modifications to a line of rods creates a line defect. In Fig. 11.1b the radius of one of the rods has been changed in comparison with the rest of the rods, and acts as a point defect. In addition, removing a line of rods, as in Fig. 11.1c, creates a line defect in a square PhC structure. The defect mode cannot propagate through the rest of the structure, because it is within the PBG spectral range. In other words, the light is trapped in this area encountering several total internal reflections. Introducing a single point defect in a photonic crystal waveguide structure results in a photonic crystal-based bandpass optical filter (Veronis *et al*., 2004). The geometry of such a structure is shown in the inset of Fig. 11.1d. The distance between adjacent rods is a, and their radius is $0.2a$. The radius of the central dielectric rod is $r_d = 0.4a$. The width of the dielectric waveguides is $0.35a$, and their distance from the center of the closest rod is $0.4a$.

Engineering the topology of the point defect can provide the required resonance frequency and quality factor (Q) of microcavities for various applications, where the cavity Q-factor is determined from the spectral width ($\Delta\lambda$) of the transmission spectrum ($Q \approx \lambda/\Delta\lambda$) for the cavity mode resonance at λ. While this is valid for a low Q-factor cavity, in general the cavity Q-factor

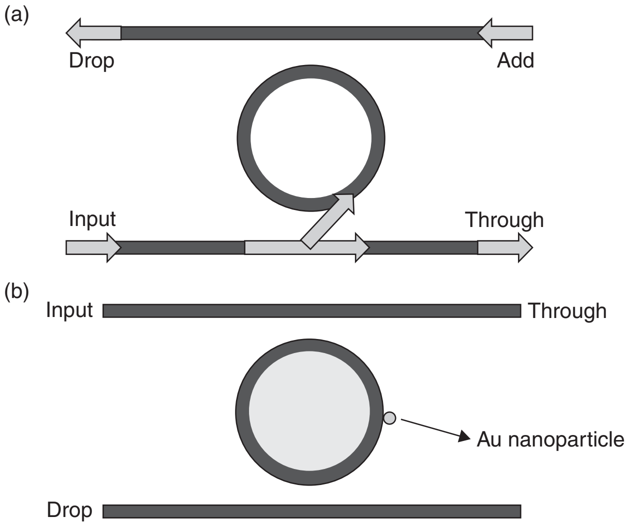


11.1 (a) Schematic of a 2-D square photonic crystal with lattice constant *a.* (b) Schematic of a point defect in a photonic crystal. (c) Schematic of a line defect in a photonic crystal. (d) Transmission spectrum of a photonic-crystal-based bandpass optical filter (Veronis *et al.*, 2004).

may be obtained by estimating the cavity photon lifetime (τ) from $Q = \omega_0 \tau$, where ω_0 is the resonant cavity mode frequency (Vahala, 2003). There have been several investigations by researchers on enhancing the Q-factors of the resonant modes. Line defects, which can guide EM waves, are used in waveguide applications and are a crucial component in photonic crystal integrated circuit applications. A critical property of photonic crystal waveguides is that the group velocity of guided modes can be greatly decreased compared to conventional waveguides leading to slow-light applications.

11.3 Ring resonators

Ring resonators are ring-shaped waveguides coupled to one or two adjacent bus waveguides. The bus waveguides serve as Input and Output ports for the ring resonator, as shown in Fig. 11.2a.



11.2 (a) Schematic of a four-port microring resonator. (b) Schematic of an on-chip four-port microring resonator configuration in which a metallic nanoparticle is adsorbed on the microring resonator (Haddadpour *et al.*, 2010).

A fraction of the incident power in the Input port couples to the ring resonator and propagates around the ring, while the rest of the power gets to the Through port (Fig. 11.2a). When the wave propagating around the ring builds up a round trip phase shift that equals an integral multiple of 2π, constructive interference takes place and the ring is on resonance. When the ring resonator is on resonance, there is a large build-up of light power in the ring. If the internal loss in the ring is negligible and the coupling between the ring and the two bus waveguides is identical, there is full transfer of power from the Input port to the Drop port, and the transmission to the Through port drops to zero (Yariv, 2000) (Fig. 11.2a). Conversely, if one wishes to add a channel to the main bus waveguide, then this channel, which is incident from the secondary bus waveguide (Add port), has to be resonant with the mode of the ring resonator. The channel is then transferred to the main bus waveguide (Through port). Thus the ring resonator coupled to two bus waveguides serves as an add/drop filter.

Various parameters such as the refractive index of the waveguide, the distance between the bus waveguides, and the optical length of the ring determine the portion of the power coupled to and from the ring resonator. In addition to the above-mentioned parameters, another crucial factor that influences the coupling strength is the lithography and patterning procedure to fabricate the gap. E-beam lithography is usually preferred in these procedures.

11.3.1 Optical filters based on ring resonators

Optical ring resonators have found many applications, and are therefore a key component for integrated optical circuits. One of the main applications of ring resonators in integrated optics is in optical filters. Optical filters are used for channel adding and dropping in optical fiber-based communication systems. As an example, an optical drop filter is used to drop a specific wavelength from a bunch of wavelengths in an optical fiber. Numerous optical structures and devices such as arrayed waveguide gratings (AWG), thin-film dielectric interference filters, and fiber Bragg gratings have been previously used as optical filters. Compared to these optical filter devices, ring resonator based optical filters have some important advantages. An optical filter must have a high quality factor Q in order to accommodate many channels. In ring resonators very high spectral selectivity can be achieved due to the high resonator quality factor Q, and the ring's intrinsic single mode nature. Recently, quality factors Q of ~760 000 have been reported (Griffith *et al.*, 2012). In addition, ring resonators also have high modal confinement and small size (radii as small as 0.9 μm are possible using a silicon/silica platform (Lipson, 2005)).

11.3.2 Active optical devices using ring resonators

In order to control the flow of light, active devices such as modulators and switches need to be developed, in addition to passive devices. In active optical devices, typically a change in the refractive index of the material induces a change in the transmission properties of the device. For example, in silicon the most effective mechanism for changing the refractive index at a fast rate is the free-carrier plasma-dispersion effect, which also has the advantage of being polarization independent (Lipson, 2005). The induced refractive index changes are associated with changes in the free-carrier electron and hole concentrations. The free-carrier concentration can be varied by injecting or generating carriers. In an electro-optic device, carriers are injected by applying an electric field to the device. In an all-optical device, carriers are generated within the device using an optical pump source (Lipson, 2005).

Due to the weak changes in refractive index, optical modulators have relatively long lengths in the order of millimeters. The device dimensions can be drastically reduced using ring resonators. When the ring resonator is on resonance, light is trapped in the resonator. When the light circulates in the ring resonator at the resonance wavelength, the optical path length increases without increasing the physical length of the device. Thus, a small change in the refractive index can significantly change the transmission near resonance, resulting in a strong intensity modulation of the signal transmitted through the ring resonator (Lipson, 2005). Ring resonators can also be used for building other compact active optical devices, such as photonic logic gates.

11.3.3 Biochemical sensors based on ring resonators

There are several different device configurations for optical biosensors. Optical ring resonators have many advantages when used as label-free biochemical sensors, including their high sensitivity, small size, and low-cost fabrication. Thus, they have recently been used extensively as a new class of optical biosensors. As in most optical sensors, the detection mechanism in ring resonator sensors is based on refractive index changes. When molecules bind to the ring resonator surface, the effective refractive index on the ring resonator surface changes. This results in a spectral shift in the sensor optical transmission response. In this case the sensitivity of the sensor is defined as the amount of spectral shift in response to a given refractive index change.

The performance of microring resonators as biochemical sensors can be improved by using labels such as metallic nanoparticles (Haddadpour and Yi, 2010), as shown in Fig. 11.2b. Adding metallic nanoparticles on the ring surface leads to the creation of 'hot spots,' where the local field, and therefore the sensitivity to refractive index variations, is greatly enhanced. Using such an approach may push the sensitivity towards single-molecule detection.

We finally note that a disadvantage of using ring resonators for modulation or sensing is the high temperature sensitivity of the device. One approach to minimize the effect of the temperature variations is the use of strain in the waveguide structure to compensate for the effects of temperature variations on the index of refraction (Lipson, 2005). Another approach is to use materials with low temperature sensitivity, such as polymers. In a ring resonator based on polymer, the sensitivity of the device to temperature is greatly reduced (Wang *et al.*, 2012). Polymer-based optical biosensors have some advantages, compared with those built on other material platforms, such as low-cost, easy fabrication, and easy surface functionalization for the detection of biological molecules.

11.4 Extraordinary optical transmission through subwavelength apertures

The existence of 100 million similar holes on the surface of a gold film was observed in the early 1990s by Thomas Ebbesen at the NEC Research Institute at Princeton, New Jersey, under an electronic microscope. Standard aperture theory predicts that the transmission efficiency of a single subwavelength cylindrical aperture scales as $(r/\lambda)^4$, where r is the aperture radius, and λ is the wavelength. Since the hole radius in the experiment was 150 nm, based on standard aperture theory less than 0.05% of the light incident on the holes could go through them for $\lambda > 1000$ nm. However, Ebbesen found that the transmission efficiency exceeded unity (when normalized to the area of the holes) for specific wavelengths. Thus, the metal film behaved as a funnel that steered the incoming light to the holes. At that time he could not find any theoretical explanation for these phenomena, and did not publish his strange results (Schechter, 2003).

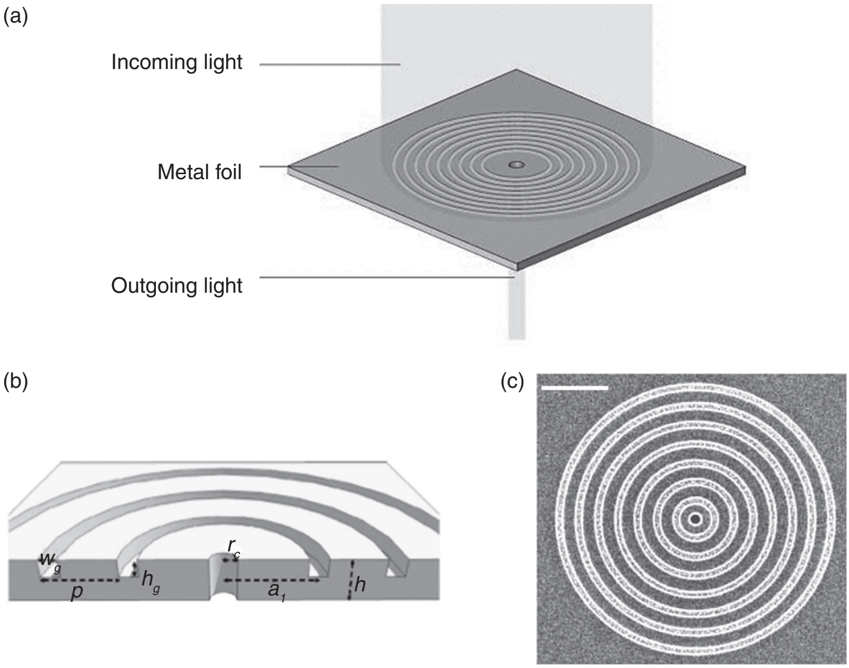
Later in 1998, Peter Wolff joined NEC and became interested in Ebbesen's old experiments. He considered more deeply the behavior of the free electrons in metals and showed that, because electrons can move freely in the surface of metals like a 2-D ocean, the incident light can interact with them. Furthermore, if the frequency of the light is the same as the resonant frequency of the electrons, it will cause the phenomena observed by Ebbesen in his experiments. Wolff showed that electrons on a metal surface excited by the incoming light can vibrate and make waves called surface plasmons (Schechter, 2003).

Wolff and Ebbesen demonstrated that when the light and surface plasmons have the same energy and momentum, they can convert to each other. The reason for more than 100% transmission of light through the subwavelength apertures in the gold film is that for specific wavelengths the incident light after hitting the surface converts to surface plasmons and directs to the holes, where it can excite new surface plasmons on the other side of the foil and convert to light again (Ebbesen *et al.*, 1998).

This observation, in addition to making the scientists reconsider their theories, has opened the new field of nanoplasmonics, dealing with the interaction of light with nanoscale metallic structures and devices, which could lead to truly novel photonic devices. Surface plasmons have been predicted since around 1900. However, it was only in recent years that fabrication of nanometallic structures made it possible to investigate the interaction of such structures with light.

Different nanostructures have been used around the apertures on the input side, such as corrugations, arrays of grooves, or gratings to excite the surface plasmons, as well as confine them and lead them through the apertures (Fig. 11.3a). In addition, corrugated metal films on the exit side of the apertures, where regenerated surface plasmons travel, can be used for beaming of light (Carretero-Palacios, 2011). By accurate design of the geometrical parameters of the grooves, the beam direction, width, and wavelength of the beaming light can be tuned.

The large number of potential applications for this enhancement of transmission led to a lot of interest in new fabrication methods and technologies with improved accuracy that could enable the realization of the designed nanostructures. A schematic of cylindrical grooves formed around



11.3 (a) Enhancement of optical transmission through a subwavelength aperture by cylindrical grooves formed around the aperture. (b) Schematic of groove arrays around a subwavelength aperture. (c) SEM image of the fabricated structure designed to enhance optical transmission at a wavelength of 660 nm. The scale bar corresponds to 2 μm.

a subwavelength aperture, and a scanning electron microscope (SEM) image of the fabricated structure, are shown in Figs 11.3b and 11.3c, respectively. Several design parameters should be considered and accurately controlled during the fabrication process: the film thickness, the depth and width of the grooves as well as the distance between consecutive grooves, the radius of the central hole, and the distance between the central hole and its nearest groove. The concentration and extraordinary transmission of light through small apertures can be used in various applications including photodetectors, optical lenses, and solar cells.

11.5 Optical nanoantennas

For many years the diffraction limit was a fundamental limitation in miniaturizing optical devices. By using metals in plasmonic devices, light confinement in subwavelength areas became possible. In recent years, using metals in nanoantennas helped to concentrate free propagating radiated energy, which holds promise for manipulation of optical fields at the nanometer scale.

While the motivation for the development of radiowave and microwave antennas was solving problems in communications, the motivation for the invention of optical antennas was microscopy. Similarly to their longer wavelength counterparts, optical antennas convert free propagating radiated energy to localized energy, and vice versa. In general, they can be used either as transmitters that radiate the localized energy, or as receivers that detect and localize the radiated energy in a specific direction (Novotny and Hulst, 2011).

The antenna efficiency $\varepsilon_{\mathrm{rad}}$ is the ratio of radiated power P_{rad} to total power P:

$$\varepsilon_{\mathrm{rad}} = \frac{P_{\mathrm{rad}}}{P} = \frac{P_{\mathrm{rad}}}{P_{\mathrm{rad}} + P_{\mathrm{loss}}} \qquad [11.1]$$

where P_{loss} is the power dissipated by absorption in the antenna or through other means.

The antenna directivity $D(\theta,\phi)$ is a measure of the antenna's ability to concentrate the energy in a specific direction:

$$D(\theta,\phi) = \frac{4\pi}{P_{\mathrm{rad}}} p(\theta,\phi) \qquad [11.2]$$

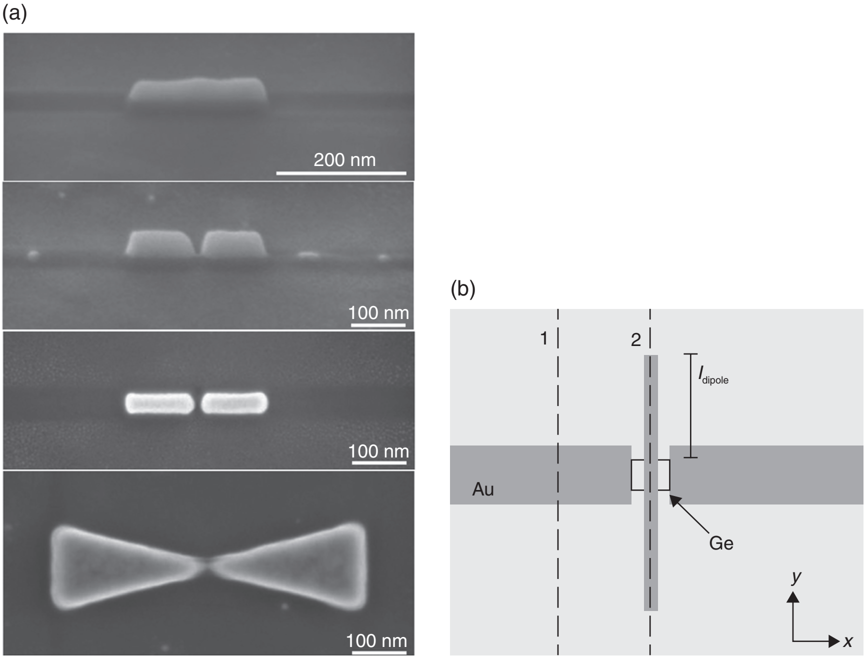
where $p(\theta,\phi)$ is the angular power density.

The antenna gain $G(\theta,\phi)$ is the product of efficiency and directivity:

$$G(\theta,\phi) = \frac{4\pi}{P} p(\theta,\phi) = \varepsilon_{\mathrm{rad}} D(\theta,\phi) \qquad [11.3]$$

For the design of conventional radiowave or microwave antennas there are rules that relate the operating wavelength to the antenna dimensions. Since at optical frequencies the light penetration in metals or skin depth is comparable to the dimensions of the antenna, these rules are not valid any more. In this case an effective wavelength can be defined, which is given by Novotny and Hulst (2011)

$$\lambda_{\mathrm{eff}} = n_1 + n_2 \left(\frac{\lambda}{\lambda_p} \right) \qquad [11.4]$$



11.4 (a) Various antennas with gap sizes down to ~ 10 nm fabricated by focused ion beam milling. (b) Top view of a photodetector consisting of a dipole antenna oriented in the y direction and two line electrodes (sleeves) in the x direction.

where n_1 and n_2 are geometric constants, and λ_p is the plasma wavelength.

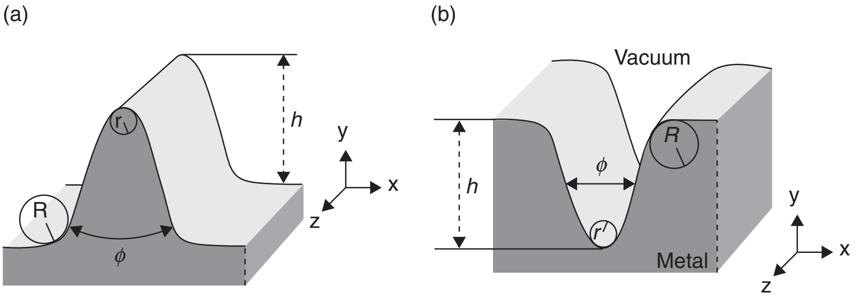
Many works and studies have been recently done on different applications of optical antennas (Fig. 11.4a). The main goal in photodetectors and photovoltaic devices is to convert light to electric current. By taking advantage of the ability of optical antennas to concentrate light into subwavelength regions, more effective detectors can be designed. For example, as shown in Fig. 11.4b, by using an optical dipole antenna near the active area of a photodetector made of germanium, a measurable photocurrent in the subwavelength region was achieved (Tang *et al.*, 2008). In addition, high concentration of light is achievable by using optical antennas shaped as nano-bowties. By placing nonlinear materials such as argon in this region it is possible to enhance nonlinear effects and generate extreme ultraviolet radiation (Schuller *et al.*, 2010). At optical frequencies metals are also highly nonlinear. This characteristic has led to many additional nonlinear optics applications of nanoantennas (Novotny and Hulst, 2011). Because of the fact that the properties of nanoparticle antennas, such as resonance frequency or scattering strength, depend on their environment, they can be used as a local probe to measure some sample properties (Bharadwaj *et al.*, 2009).

Many other applications for optical antennas have been studied and hold promise in the future. For example, recently for the first time a bright molecule emission with a lifetime of ~20 ps (super-bright single-photon) was observed in the gap of a bow-tie antenna, which could be useful in quantum optics applications. Although many of these antennas have been fabricated, there are still several challenges which should be considered. First, as mentioned above, the rules used for the design of conventional radiowave or microwave antennas do not apply to optical nanoantennas. Thus, new antenna concepts and designs are required. In addition, impedance matching is a major challenge for optical nanoantennas, just as with any kind of antenna.

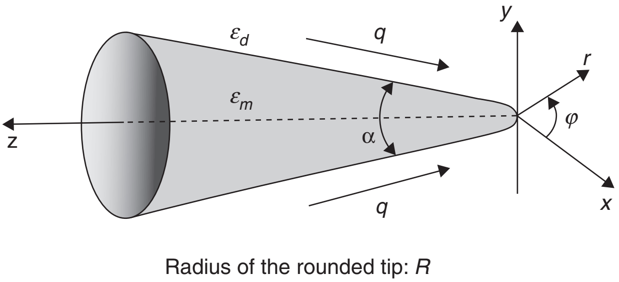
11.6 Plasmonic focusing

Focusing and localization of light is an interesting research topic, since there are many applications that need high concentration and high resolution of light, such as optical sensors, beaming, and imaging in optical microscopes. For many years the major obstacle in achieving this goal was the diffraction limit, which does not allow the concentration of EM waves in dielectric media in a region much smaller than the wavelength. Surface plasmons and the propagation of light along the interface between a material with a negative real part of the dielectric constant (such as metals), and a material with a positive real part of the dielectric constant (such as dielectrics) could address this problem. The field enhancement and confinement of light in a small region is the most significant characteristic of plasmonic devices. However, it should be pointed out that there is a tradeoff between the light confinement and the propagation length of waveguide modes, since higher concentration results in increased losses. Although higher energy confinement is accompanied by higher dissipation of light, proper design of plasmonic devices can lead to substantial intensity enhancements that offset the power dissipation (Choi *et al.*, 2009). There are different types of nanometallic waveguides that can guide light in subwavelength regions, including V-groove (Fig. 11.5a) and wedge waveguides (Fig. 11.5b). In these structures the light can be highly localized in the two dimensions (x and y), and propagate in the third dimension (z) (Yan and Qiu, 2007).

When such structures are tapered, they can be used for nanofocusing. It has been demonstrated that high field enhancements can be obtained near the tip of such tapered metallic nanostructures. Various geometries were proposed and investigated theoretically and experimentally in recent years for plasmon focusing, such as tapered metallic rods, sharp metallic wedges, and tapered nanogaps between two metallic media (Nerkarayan,



11.5 Schematics of (a) V-groove and (b) wedge plasmonic waveguides.



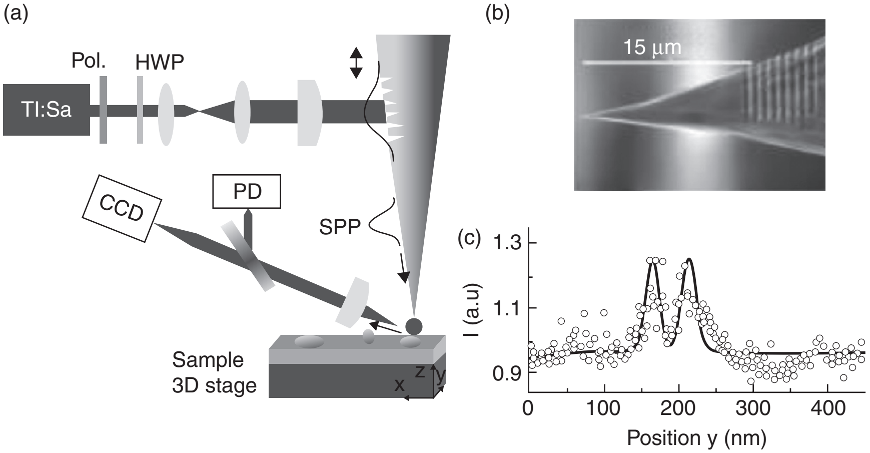
11.6 Schematic of a metallic rod with taper angle α. The permittivities of the metal (rod) and surrounding dielectric are ε_m and ε_d, respectively, and q is the wave vector of the surface plasmon propagating towards the tip of the rod.

1997; Stockman, 2004). A schematic of a tapered metallic rod is shown in Fig. 11.6. There are three key geometrical parameters that should be considered for practical applications and fabrication: the angle of the taper, the curvature of the tip, and the length of the taper. Each of these three parameters can affect the enhancement of light and also the energy confinement (Gramotnev *et al.*, 2008).

Depending on the angle of the tip, the focusing can happen in the adiabatic or non-adiabatic regime. If the taper angle is small enough so that variation of the plasmon wave number within one wavelength distance from the tip is negligible, then the adiabatic or geometrical optics approximation can be used. For example, it was shown that this approximation is valid up to an angle of 35° for a gold tapered rod in vacuum at a wavelength of 632.8 nm. If the angle of the taper is larger than this, then variation of the plasmon wave number within one wavelength distance from the tip is non-negligible, and this is referred to as the non-adiabatic regime. In this case there is an optimal length and also an optimal angle for achieving maximum enhancement of energy. For the optimized tapered metallic rod nanofocusing structure, a field enhancement of ~2000 is achievable in regions of a few square nanometers (Gramotnev and Bozhevolnyi, 2010; Gramotnev *et al.*, 2008).

The other critical parameter that should be considered in fabrication of tapered metallic rods is the curvature of the tip. For a gold tapered rod with an angle of ~35° and length of ~10 μm, which are optimal at a wavelength of 632.8 nm, it was shown that decreasing the tip radius from 10 to 2 nm improves the field enhancement from ~200 to ~1600 (Gramotnev *et al.*, 2008). So far the strongest enhancement was predicted to be in tapered metallic rods. The localization in a single wedge is smaller, about 10–20 nm. Although other structures, such as tapered nanogaps, can also lead to high confinements in nanometer scale regions, their field enhancement is up to ~100 with a tip of a few nanometers radius (Gramotnev and Bozhevolnyi, 2010).

Plasmonic nanofocusing is currently used in various applications. Scattering near-field scanning optical microscopy (s-NSOM) is a technique that can use plasmonic focusing to achieve very high resolution for imaging. As shown in Fig. 11.7, sub-30 nm resolution imaging of metallic nanoparticles



11.7 (a) Schematic of an adiabatic nanofocusing scattering-type optical microscope (s-NSOM). Light from a tunable Ti:sapphire laser at 780 nm passes through a polarizer (Pol.) and a half wave plate (HWP), and is focused by a high numerical aperture objective onto a nanoslit grating milled onto the shaft of a gold taper. Light scattered from the tip apex into the far field is collected by a second high numerical aperture objective in reflection geometry and imaged onto a photodetector. (b) Scanning electron micrograph of a conical gold tip with a 780 nm grating period prepared by focused ion beam milling. (c) Cross-sections of the optical intensity along the y direction. The strong near-field enhancement at the edges of the short axis of the elliptical gold nanoparticle (with 100 × 40 × 15 nm^3 dimensions) indicates that the component of the local electric field oriented along the tip axis (z-direction) is imaged.

was recently reported using surface plasmons launched onto smooth gold tapers where they were adiabatically focused towards the nanometer-sized taper apex (Sadiq *et al.*, 2011).

11.7 Near-field optical microscopy

By using near-field scanning optical microscopy (NSOM), the diffraction limit of light in optical microscopes can be circumvented, and very high resolution can be achieved. In this section we discuss aperture and apertureless NSOM designs, as well as a new NSOM technique that combines these two designs.

11.7.1 Aperture probe near-field scanning optical microscopy (NSOM)

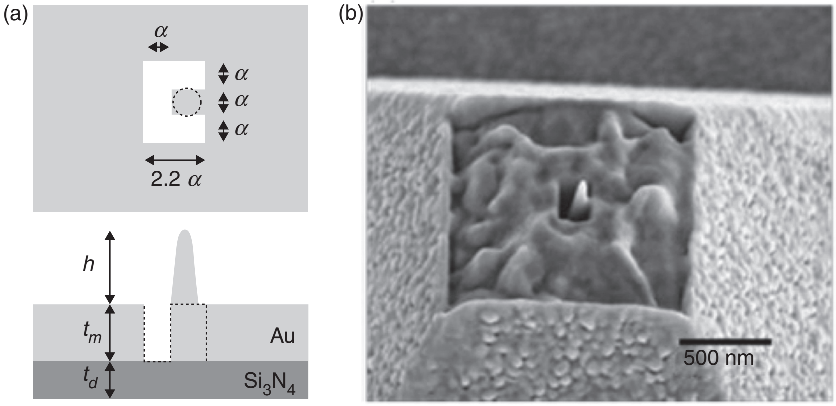
One approach for near-field optical microscopy is based on using subwavelength apertures in metal-coated optical probes, and scanning a few nanometers above the sample. In such microscopes, the aperture size rather than the wavelength of light determines the resolution. However, decreasing the aperture size will also greatly decrease the optical throughput. As discussed in Section 11.4, for subwavelength apertures the power throughput, defined as the ratio of the total transmitted power to the incident power over the aperture area, scales as $(d/\lambda)^4$, where d is the aperture size (Shi and Hesselink, 2002). Thus, the aperture size is a limitation for the performance of these systems. Choosing an appropriate shape design for the apertures, such as C-shaped or bow-tie, can lead to substantial enhancements in the power throughput. For C-shaped apertures the light is localized on the ridge at the exit surface of the aperture, and provides a very strong near-field. Using such structures, an enhancement of power throughput of up to two or three orders of magnitude higher than square apertures is achievable. However, in practice the optical spot sizes for such C-shaped structures are around 40–70 nm, limited by present fabrication technology (Cheng *et al.*, 2011).

11.7.2 Apertureless probe NSOM

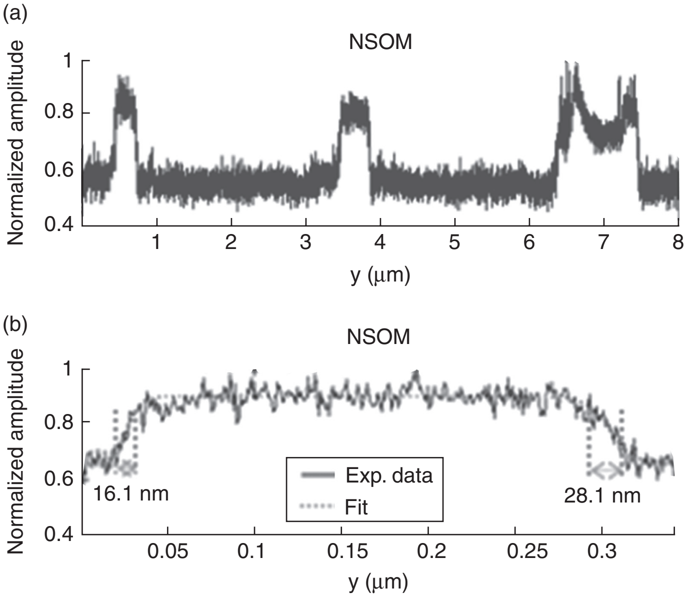
An alternative technique to obtain higher resolution is based on using an apertureless probe in near-field scanning optical microscopy (aNSOM). By using a sharp metallic nanotip, a very high concentration of light can be achieved. High resolution, with optical spot size of less than 10 nm, was reported using apertureless probes, due to the sharpness of the tip apex, which determines the optical spot size (Gerton *et al.*, 2004). The problem that appears here is the background illumination. The light coming from the laser light source also illuminates a large area around the tip apex, and interferes with the near-field scattered from the sample. In spite of the very high concentration and enhancement of energy on the tip of the apertureless probe NSOM, background optical energy always exists, and therefore applying a modulation technique is necessary. When using such a technique, the detector is able to detect only the near-field contribution, which is a modulated signal, and reject the unmodulated background light (Schuller *et al.*, 2010).

11.7.3 CAN-tip probe NSOM

By taking advantage of the characteristics of both aperture probes for blocking the background signal, and apertureless probes for high energy concentration and resolution, a new NSOM technique combining these two was first proposed by Frey *et al.* (2002). A metallic tip formed on the aperture of a conventional fiber probe, which is referred to as tip-on-aperture probe, was introduced. In such a probe, the light passing through the aperture is concentrated by a sharp nanotip attached to the aperture. In a more recent NSOM experiment, a resolution of 16.1 nm was obtained by using a C-shaped aperture combined with a nanotip (CAN-Tip). The nanotip, which acts as a metallic nanoantenna, is extended from the ridge of the C-shaped metallic nanoaperture (Cheng *et al.*, 2011). When operating at the resonant frequency, the charges concentrate at the corners of the ridge on the exit surface of the C-shaped aperture. In addition, surface charges concentrate on the tip of the nanoantenna due to the lightning-rod effect (electric fields are strongest at the sharpest features of the object), and make a very-high-intensity ultra-small optical spot. Figure 11.8 shows the schematic and SEM image of a fabricated CAN-Tip. By designing the aperture and the nanoantenna



11.8 (a) Schematic design of the CAN-Tip. (b) 35° angled view SEM images of a CAN-Tip fabricated with focused ion beam milling.



11.9 (a) NSOM response when scanning across three Cr nanodisks with a CAN-Tip NSOM probe set different distances away from the center of the disks. The signal-to-noise ratio (SNR) is ~8.9. (b) Close-up of the first disk scan from (a) (solid line). The overlaid dashed line shows the fitted data. The narrowest transition at the left edge shows an equivalent 16.1 nm full width at half maximum (FWHM) Gaussian transition. This demonstrates the 16.1 nm optical resolution of the CAN-Tip NSOM probe.

to resonate at a wavelength of 980 nm, Cheng *et al.* (2011) predicted high intensity (650×) and high optical resolution (~$\lambda/60$) for such a structure. In addition, no modulation technique is required due to the absence of background signal. Figure 11.9 shows the CAN-Tip NSOM response when scanning across three Cr nanodisks with a probe set at different distances from the center of the disks. Optical resolution of 16.1 nm was demonstrated in this experiment.

CAN-Tip could potentially be used in several other near-field technologies that require high optical intensity with nanoscale resolution, such as near-field optical recording, heat-assisted magnetic recording (HAMR), tip-enhanced Raman scattering (TERS), and nanoscale electron source generation (Cheng *et al.*, 2011).

11.8 Plasmonic waveguides

As mentioned in Section 11.6, several different plasmonic waveguiding structures have been proposed to guide subwavelength optical modes, such as metallic nanowires and metallic nanoparticle arrays. Most of these structures support a highly-confined mode only near the surface plasmon frequency. In this regime, the optical mode typically has low group velocity and short propagation length. It has been shown, however, that a metal–dielectric–metal (MDM) structure with a dielectric region thickness of ~100 nm supports a propagating mode with a nanoscale modal size at a wavelength range extending from DC to visible. Thus, such a waveguide could be potentially important in providing an interface between conventional optics and subwavelength electronic and optoelectronic devices. Because of the predicted attractive properties of MDM waveguides, their modal structure has been studied in great detail, and their structures have also been experimentally explored (Dionne *et al.*, 2006). Recent research work has focused on the development of functional plasmonic devices, including active devices, for nanoscale plasmonic integrated circuits.

The performance of bends and splitters in MDM subwavelength plasmonic waveguides was investigated (Veronis and Fan, 2005). It was shown that bends and splitters with no additional loss over a very wide frequency range can be designed for MDM waveguides with center layer thickness small compared to the wavelength. The concept of characteristic impedance used in microwave engineering was introduced for such systems to account for their behavior. In addition, compact couplers between high-index contrast dielectric slab waveguides and MDM subwavelength plasmonic waveguides have been investigated. Couplers created by simply placing a dielectric waveguide terminated flat at the exit end of a plasmonic waveguide can be designed to have high transmission efficiencies at the optical communication wavelength. The transmission efficiency of the couplers can be further increased by using optimized multisection tapers. In both cases the transmission response is broadband.

Switches for MDM plasmonic waveguides have also been introduced. An MDM waveguide directly coupled to a cavity filled with an active material with tunable absorption coefficient can act as an absorption switch, in which the on/off states correspond to the absence/presence of optical pumping. In addition, an MDM plasmonic waveguide side-coupled to a cavity filled with an active material can operate as an absorption switch, in which the on/off states correspond to the presence/absence of pumping. For a specific modulation depth, the side-coupled-cavity switch results in more compact designs compared to the direct-coupled-cavity switch.

11.9 Enhancement of nonlinear processes

Nonlinearity in optics is introduced by alterations in the optical properties of a medium as the intensity of the applied field is increased, or when one or more other fields are introduced. Nonlinear optical effects include: electrooptical effects (Pockels and Kerr effects); the Zeeman effect, in which spectral lines are split by a magnetic field; and the magneto-optical effect, which is magnetic-field-induced double refraction (Cotton-Mouton and Faraday effects). Nonlinear optical processes have numerous applications in optics. These include wavelength conversion, field-dependent propagation, optical limiters and multistable cavities, and nonlinear signal propagation. However, the nonlinear coefficients of naturally occurring optical materials are not sufficiently large and, hence, a large interaction volume or high excitation power is required to generate acceptable levels of nonlinear signals. If the effective nonlinearity of existing materials could be increased, there would be numerous benefits. The small absolute scale of the nonlinear coefficients and phase mismatch effects are two significant factors that restrict the efficiency of the nonlinear processes. Nanophotonic structures and devices may offer a means to compensate for these limitations.

Second-harmonic generation (SHG) is the second order nonlinear process for which two photons, with the same frequency ω, interact simultaneously with matter to generate a photon of frequency 2ω. In recent years the generation of second-harmonic radiation at metal surfaces has been studied extensively, because the intensity of the produced harmonic can be strongly enhanced due to localized surface plasmons. Similarly, metal nanoparticles arranged in ordered patterns give rise to a nonlinear response due to particle plasmons, which can be enhanced by introducing asymmetry in the local field distribution (Canfield *et al.*, 2007). SHG in a set of nanoparticles with proper asymmetry, and plasmonic solitons in metal–dielectric structures, have been considered by Zheludev and coworkers (Zheludev and Emelyanov, 2004) and Samson and coworkers (Samson *et al.*, 2011), respectively.

Surface enhanced Raman scattering (SERS) is a technique that may result in orders of magnitude increase in the Raman signal intensity, as high as 10^{14}–10^{15}. SERS is of interest for trace material analysis, flow cytometry, and other applications where the current sensitivity and/or speed of a Raman measurement are insufficient. The enhancement takes place at a metal surface, which has nanoscale roughness. SERS occurs when molecules are brought to the rough metal surface. The enhancement in the Raman signal is partially attributed to the large local electric field enhancement in specific locations of the metal film known as ‘hot spots.’ Smooth surfaces do not result in enhancement. Large enhancements are observed for silver, gold, and copper surfaces. In addition to rough metal films, several other geometries can be used for SERS, such as metal films structured with hole arrays, metallic nanoparticles, as well as metallic nanowire arrays (Maier, 2007).

Nonlinear optical processes can also be enhanced with PhCs. Nonlinear photonic crystal microresonators can enhance a variety of nonlinear optical processes. It has been demonstrated that PhC microresonators can enhance the performance of traditional nonlinear devices, such as optical switches and nonlinear Fabry-Pérot interferometers, by several orders of magnitude. In addition, due to their design versatility, PhC cavities can be used as the basis of completely new configurations performing all-optical functions, such as all-optical transistors (Yanik *et al*., 2003). Moreover, PhC resonators offer new fundamental ways of tailoring optical nonlinearities by using the so-called Purcell effect. Finally, in addition to strong enhancement of nonlinear phenomena through PhC resonators, nonlinear effects can also be enhanced by using the slow-light properties of PhCs, via the corresponding band edges (D'Aguanno *et al*., 2001) or by means of coupled-cavity waveguides (Soljacic *et al*., 2002). Some of the nonlinear optical phenomena that can be greatly enhanced by PhC resonators include harmonic generation and optical bistability (Bravo-Abad *et al*., 2007).

11.10 Application in photovoltaics

Solar cells can provide virtually unlimited amounts of energy by effectively converting sunlight into clean electrical power. Large-scale implementation of photovoltaic technology hinges on our ability to inexpensively produce high-efficiency solar cells (Schuller *et al*., 2010). Thin-film solar cells could provide a viable pathway towards this goal by offering low materials and processing costs. However, the energy conversion efficiencies of such cells are fairly low, owing to the large mismatch between electronic and photonic length scales in these devices. For photon energies close to the bandgap, the absorption length of light in semiconductors is significantly longer than the electronic diffusion length in most deposited thin-film materials (Schuller *et al*., 2010). In other words, the thickness of the active region of these devices is limited to less than the electronic diffusion length. This, however, limits the photon absorption, since the light absorption length is significantly longer than the active region thickness.

Plasmonic nanostructures may provide a strategy to solve the issue above by dramatically improving light absorption in thin-film photovoltaic cells. Present research on using plasmonics to boost photovoltaic cell efficiencies is focused on taking advantage of the high near-fields surrounding metallic nanostructures close to their surface plasmon resonance frequency, and effective light trapping in the active semiconductor layer to boost light absorption (Schuller *et al*., 2010).

11.11 Conclusion

Nanophotonics is the science and engineering of light-matter interactions at the nanoscale. In this chapter we first discussed wavelength-scale nanophotonic devices, such as PhCs and ring resonators. We then discussed several topics associated with manipulating light at subwavelength scales with nanometallic structures: extraordinary optical transmission through subwavelength apertures, optical nanoantennas, plasmonic focusing, near-field optical microscopy, and plasmonic waveguides. Finally, we discussed the application of nanophotonics in enhancement of nonlinear processes and photovoltaics.

11.12 References

Bharadwaj P, Deutsch B and Novotny L (2009), 'Optical antennas', *Adv Opt Photon*, **1**, 438–483.

Bravo-Abad J, Rodriguez A, Bermel P, Johnson S G, Joannopoulos J D and Soljacic M (2007), 'Enhanced nonlinear optics in photonic-crystal Microcavities', *Opt Express*, **15**, 16161.

Canfield B K, Husu H, Laukkanen J, Bai B, Kuittinen M, Turunen J and Kauranen M (2007), 'Local field asymmetry drives second-harmonic generation in noncentrosymmetric nanodimers', *Nano Lett*, **7**, 1251–1255.

Carretero-Palacios S, Mahboub O, Garcia-Vidal F J, Martin-Moreno L, Rodrigo S G, Genet C and Ebbesen T W (2011), 'Mechanisms for extraordinary optical transmission through bull's eye structures', *Opt Express*, **19**, 10429–10442.

Cheng Y T, TakashimaY, Yuen Y, Hansen P C, Leen J B and Hesselink L (2011), 'Ultra-high resolution resonant C-shaped aperture nano-tip', *Opt Express*, **19**, 5077–5085.

Choi H, Pile D F P, Nam S, Bartal G and Zhang X (2009), 'Compressing surface plasmons for nano-scale optical focusing', *Opt Express*, **17**, 7519–7524.

D'Aguanno G, Centini M, Scalora M, Dumeige Y, Vidakovic P, Levenson J A, Bloemer M J, Bowden C M, Haus J W and Bertolotti M (2001), 'Photonic band edge effects in finite structures and applications to $\chi(2)$ interactions', *Phys Rev E*, **64**, 016609.

Dionne J A, Lezec H J and Atwater H A (2006), 'Highly confined photon transport in subwavelength metallic slot waveguides', *Nano Lett*, **6**, 1928–1932.

Ebbesen T W, Lezec H J, Ghaemi H F, Thio T and Wolff P A (1998), 'Extraordinary optical transmission through sub-wavelength hole arrays', *Nature*, **391**, 667–669.

Frey H G, Keilmann F, Kriele A and Guckenberger R (2002), 'Enhancing the resolution of scanning near-field optical microscopy by a metal tip grown on an aperture probe', *Appl Phys Lett*, **81**, 5030–5032.

Gerton J M, Wade L A, Lessard G A, Ma Z and Quake S R (2004), 'Tip-enhanced fluorescence microscopy at 10 nanometer resolution', *Phys Rev Lett*, **93**, 180801.

Gramotnev D K, Vogel M W and Stockman M I (2008), 'Optimized nonadiabatic nanofocusing of plasmons by tapered metal rods', *J Appl Phys*, **104**, 034311.

Gramotnev D K and Bozhevolnyi S I (2010), 'Plasmonics beyond the diffraction limit', *Nat Photonics*, **4**, 83–91.

Griffith A, Cardenas J, Poitras C B and Lipson M (2012), 'High quality factor and high confinement silicon resonators using etchless process', *Opt Express*, **20**, 21341–21345.

Haddadpour A and Yi Y (2010), 'Metallic nanoparticle on micro ring resonator for bio optical detection and sensing', *Biomed Opt Express*, **1**, 378–384.

Lipson M (2005), 'Guiding, modulating, and emitting light on silicon – challenges and opportunities', *J Lightwave Technol*, **23**, 4222–4238.

Maier S A, Kik P G, Atwater H A, Meltzer S, Harel E, Koel B E and Requicha A A G (2003), 'Local detection of electromagnetic energy transport below the diffraction limit in metal nanoparticle plasmon waveguides', *Nat Mater*, **2**, 229–232.

Maier S A (2007), *Plasmonics: Fundamentals and Applications*, New York, Springer.

Nerkarayan K (1997), 'Superfocusing of a surface polariton in a wedge-like structure', *Phys Lett A*, **237**, 103–105.

Novotny L and Hulst N V (2011), 'Antennas for light', *Nat Photonics*, **5**, 83–90.

Proietti Zaccaria R, Gopalakrishnan A, Das G, Gentile F, Haddadpour A, Toma A, Angelis F D, Liberale C, Mecarini F, Razzari L, Giugni A, Krahne R and Fabrizio E D (2012), 'Photonic Crystals for Plasmonics: From Fundamentals to Superhydrophobic Devices', in Massaro A. ed., *Photonic Crystals – Innovative Systems, Lasers and Waveguides*, Vienna, InTech.

Sadiq D, Shirdel J, Lee J S, Selishcheva E, Park N and Lienau C (2011), 'Adiabatic nanofocusing scattering-type optical nanoscopy of individual gold nanoparticles', *Nano Lett*, **11**, 1609–1613.

Samson Z I, Horak P, MacDonald K F and Zheludev N I (2011), 'Femtosecond surface plasmon pulse propagation', *Opt Lett*, **36**, 250–252.

Schechter B (2003), 'Bright new world', *New Sci*, **178**, 31–33.

Schuller J A, Barnard E S, Cai W, Jun Y C, White J S and Brongersma M L (2010), 'Plasmonics for extreme light concentration and manipulation', *Nat Mater*, **9**, 193–204.

Shi X and Hesselink L (2002), 'Mechanisms for enhancing power throughput from planar nanoapertures for nearfield optical data storage', *Jpn J Appl Phys*, **41**, 1632–1635.

Soljacic M, Johnson S G, Fan S, Ibanescu M, Ippen E and Joannopoulos J D (2002), 'Photonic-crystal slow-light enhancement of non-linear phase sensitivity', *J Opt Soc Am B*, **19**, 2052–2059.

Stockman M I (2004), 'Nanofocusing of optical energy in tapered plasmonic waveguides', *Phys Rev Lett*, **93**, 137404.

Tang L, Kocabas S E, Latif S, Okyay A K, Ly-Gagnon D, Saraswat K C and Miller D A (2008), 'Nanometre-scale germanium photodetector enhanced by a near-infrared dipole antenna', *Nat Photon*, **2**, 226–229.

Vahala K J (2003), 'Optical microcavities', *Nature*, **424**, 839–846.

Veronis G, Dutton RW and Fan S (2004), 'Method for sensitivity analysis of photonic crystal devices', *Opt Lett*, **29**, 2288–2290.

Veronis G and Fan S (2005), 'Bends and splitters in subwavelength metal-dielectric-metal plasmonic waveguides', *Appl Phys Lett*, **87**, 131102.

Wang L, Han X, Gu Y, Lv H, Cheng J, Teng J, Ren J, Wang J, Jian X and Zhao M (2012), 'Optical biosensors utilizing polymer-based athermal microring resonators', *Proc SPIE*, **8427**, 842731.

Yan M and Qiu M (2007), 'Guided plasmon polariton at 2D metal corners', *J Opt Soc Am B*, **24**, 2333–2342.

Yanik M F, Fan S, Soljacic M and Joannopoulos J D (2003), 'All-optical transistor action with bistable switching in a photonic crystal cross-waveguide geometry', *Opt Lett*, **28**, 2506–2508.

Yariv A (2000), 'Universal relations for coupling of optical power between microresonators and dielectric waveguides', *Elect Lett*, **36**, 321–322.

Zheludev N I and Emelyanov V I (2004), 'Phase matched second harmonic generation from nanostructured metallic surfaces', *J Opt A: Pure Appl Opt*, **6**, 6–28.

12

Nanodevices: fabrication, prospects for low dimensional devices and applications

T. DANIELS-RACE, Louisiana State University, USA

DOI: 10.1533/9780857098757.399

Abstract: In the nano-realm, the line between where a 'structure' ends and a 'device' begins can at times be a thin one. More often than not, device *characteristics* may serve as the deciding boundary, albeit with the expected arguable interdisciplinary definitional differences. Nevertheless, taking its cues from the monumental breakthroughs in terms of devices and circuits, as these have come to be defined in the field of microelectronics, this chapter provides a look into the progress toward this same level of 3D functionality for nanoscale electronic devices. Beginning with a brief overview of the historical steps that have led to modern-day nanodevice technology to select milestones illustrative of device operation, the narrative to follow is designed to offer a template for and a guide toward a more detailed study into this quite extensive topic. However, with allowance for a certain degree of author's choice as to the devices themselves, this chapter pays homage to key discoveries in the field – of both device components and the equipment used to create or characterize them – while describing some of the structures that have come to be state-of-the-art illustrations of nanodevices in this wide-ranging, dynamic, and continually evolving field.

Key words: nanodevices, single-electron transistors, nanowires, nanoarrays.

12.1 Introduction

'Make it work!' This catch phrase of a popular cable TV show would no less aptly describe the goal of device development in the nano-realm. Although 'nano' serves as the ubiquitous prefix to a wide range of technical (and non-technical) endeavors, this chapter will embrace the 1–100 nm range in terms of at least one dimension of the devices to be discussed. That being said, the umbrella of nanodevices encompasses an unusually wide range of technologies as far as 'devices' are concerned. One might consider work ranging from solid-state electronics to colloidal suspensions of nanoparticles with, in turn, applications ranging from traditional electronic circuits to drug delivery via biomedical implants. Thus, the 'history' to follow is indeed abbreviated, as the nanodevices of today can be associated with practically any application whose technological impact depends upon metrics such as speed (i.e. computing), size (i.e. consumer products, iPod © vs iPod Nano©), or even sensitivity (i.e. biomarkers for cancer detection). With such far-reaching implications, one point of 'agreement' with respect to nanodevices is that among engineers, chemists, materials scientists, physicists, and biologists – and even this list of professions is not exhaustive in its description of 'nano' researchers! – there is only *minimal agreement* as to which nanostructures are significant, or even worthy of the designation 'device.' Therefore, in an attempt to temporarily tame this unwieldy beast of a topic, we will discuss a select few milestones, so as to give you, the reader, a plausible overview and a springboard from which to launch your own more directed, or more comprehensive, study of nanodevices.

We begin with the work of the tenth century AD European artisans typically presented as one of the earliest illustrations of nanotechnology. The aforementioned used the natural changes in bulk vs nanoparticle sized gold embedded in glass to attain the various colors used to create stained glass windows for medieval churches (Ratner and Ratner, 2003). However, this process traces back to techniques used in the decoration of pottery in ninth century Mesopotamia (Binns, 2010). Moreover, archeological findings from Africa have revealed 4000-year-old techniques in lead based chemistry which, preserved through the ages, became the basis for the production of hair dyes including galena, or PbS nanocrystals. The latter have been identified in cosmetology recipes dating from the Greco-Roman period to the present day (Walter *et al.*, 1999; Walter *et al.*, 2006). Finally, the development of 'India ink,' presumed to have been invented in China around 2700 BC, involved a formulation consisting of what would today be described as carbon nanoparticles (Binns, 2010). In any case, and with all due respect to our nanotechnology forbearers, it is reasonable to assume that an understanding of the science behind these results was limited to the relative skill of the given artist or practitioner working at the time.

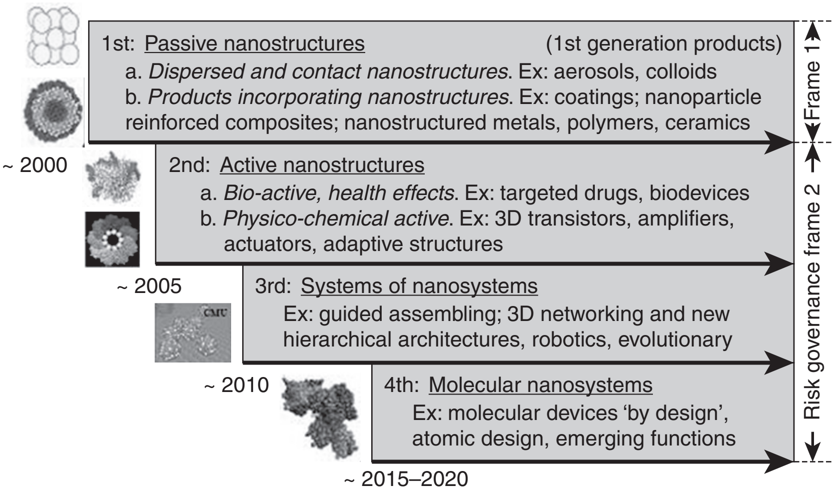
It was not until the 1930s that modern-day scientists and engineers began to seriously consider the nano-realm via the presentation of the first transmission electron microscope (TEM) by Ernst Ruska and Max Knoll. (http://www.microscopy.ethz.ch/history.htm). This breakthrough technology allowed scientists to move beyond the limitations of light microscopes (~200 nm at the time). By 1959, the now famous talk, 'There's Plenty of Room at the Bottom,' was delivered by Professor (and later Nobel laureate) Richard P. Feynman of the California Institute of Technology (Feynman, 1960). Feynman, a theoretical physicist, described the potential for future research involving the manipulation of atoms and molecules to produce and control more macroscopic functionalities. Often mentioned as a starting point for nanotechnology, in point of fact the scientific community did not really take notice of this one out of hundreds of Feynman talks until over thirty years later, when K. Eric Drexler published *The Engines of Creation: The Coming Era of Nanotechnology* (Drexler, 1992). Drexler spoke to the essence of nanotechnology from the perspectives of both ‘bottom-up’ and ‘top-down’ technologies. With respect to the former, he discussed manufacturing beginning from the atomic and molecular levels. With respect to the latter he addressed issues of magnitudes and scaling and the considerations necessary to move beyond models based on the classical continuum. Nevertheless, the hope from which to truly consider the things of which Feynman and others had spoken did not begin to realistically emerge until the introduction of another breakthrough in 1981 from Gerd Binnig and Heinrich Rohrer of IBM Zurich (Binnig *et al.*, 1982b). What came to be known as the scanning tunneling microscope (STM) earned these two researchers the 1986 Nobel Prize in Physics and opened the door to the nano-realm beyond that of prophetic predictions.

Finally, this brief history would be remiss without the mention of one more nanotechnology milestone and source of the 1996 Nobel Prize in Physics shared by Harry Kroto, Robert Curl, and the freshman chemistry teacher of this chapter’s author, Richard Smalley. These researchers are credited with the discovery of a new form of carbon (C_{60}) whose structural arrangement mimicked that of the geodesic dome design of American architect Buckminster Fuller. ‘Buckyballs’ or fullerenes, as this discovery came to be known, have led to the study of carbon nanotubes (CNT), for which researchers are claiming and finding potential applications in areas ranging from consumer electronics to medicine (Tomanek *et al.*, 2008; Gomez-Gualdron *et al.*, 2011; Ahn and Je, 2012). Even construction engineers are considering the use of CNTs in cement and other materials, as this form of carbon has exhibited tensile strengths greater than steel (Wille and Loh, 2010).

Yet amidst all of these breakthroughs and materials in the nano-realm, where, one may ask, are the *devices*, and how are they defined? The rest of this chapter will attempt to answer these questions. Again, the ubiquitous nature of the field of nanotechnology automatically warrants some level of ‘author’s choice’ in terms of devices described and discussed. However, it is our hope that this chapter will not only inform, but will inspire the reader to further investigate this groundbreaking ‘realm of the small’ as it affects the biggest of phenomena and functions of our modern technological lives.

12.2 Motivation for nanodevices

Hopefully, the previous section has set the stage to answer these questions and maybe, in some sense, already has. Yet, as an introduction to nanodevices, the question of even ‘Why Bother?’ must be addressed. Therefore, a look at the motivation for this field is in order.


1st: Passive nanostructures
(1st generation products)
a. Dispersed and contact nanostructures. Ex: aerosols, colloids
b. Products incorporating nanostructures. Ex: coatings; nanoparticle reinforced composites; nanostructured metals, polymers, ceramics
~ 2000
2nd: Active nanostructures
a. Bio-active, health effects. Ex: targeted drugs, biodevices
b. Physico-chemical active. Ex: 3D transistors, amplifiers, actuators, adaptive structures
~ 2005
3rd: Systems of nanosystems
Ex: guided assembling; 3D networking and new hierarchical architectures, robotics, evolutionary
~ 2010
4th: Molecular nanosystems
Ex: molecular devices 'by design', atomic design, emerging functions
~ 2015–2020
Frame 1
Risk governance frame 2

12.1 A representation of nanodevice development over the past decade and beyond as suggested by Renn and Roco to the International Risk Governance Council in 2006. (*Source*: Reproduced with permission from M. Rocco.)

Since the introduction of *micro*electronics into the technology landscape of the 1970s, the quest for higher speed, greater efficiency, better performance and, of course, smaller dimensions has forged ahead at a practically exponential rate. Even the original 'Moore's Law' of doubling the number of devices on an integrated circuit chip every year (or every 18 months as has also been quoted) has proven to be almost eerily accurate (Moore, 1998; Moore, 2005). Thus, as device densities increase, so must device size decrease while maintaining the metrics of speed and reliable performance. To begin to meet this challenge, electronic device engineers have defined two primary approaches – 'top-down' and 'bottom-up' – fabrication techniques. This chapter will present devices inclusive of both. In turn, researchers still debate as to whether or not the term *nanotechnology* applies to measurements and observations in the 1–100 nm range, as opposed to the actual dimensions of a given structure. Yet, insomuch as this chapter deals with devices, we will focus on the latter of the two. Finally, categorizing devices is a topic of likewise vigorous discussion. To provide a point of some consensus, we will rely on Fig. 12.1 which emerged as the result of discussions initiated by Ortwin Renn's and Mihail (Mike) Roco's 'Whitepaper no. 2' presented to the 2006 International Risk Governance Council in Geneva, Switzerland (Renn and Roco, 2006).

As shown, nanotechnological development can be divided into four generations as follows:

- **Passive nanostructures** – essentially materials designed for a single task
- **Active nanostructures** – devices designed for multi-tasking, i.e. sensors, actuators, drug delivery
- **Systems of nanosystems** – thousands of interacting devices and systems of devices
- **Molecular nanosystems** – described by the 'control and restructuring of matter at the nanoscale' or the 'engineering of functional systems at the molecular scale'

Clearly, these categories help to delineate functions, as opposed to the devices themselves. However, in the interest of contributing to this discussion of 'Why smaller? Why better?' Fig. 12.1 is noteworthy in that it provides a timeline, not so much in terms of dates, but for the various stages of what has become nanoscale device development, particularly from the previously mentioned bottom-up perspective. Essentially mimicking the intention of Moore's Law as far as decreasing size without sacrificing functionality, nanodevices represent an *opportunity* through which both dimensions and the phenomenological behavior of materials and, ultimately, devices may be best exploited to achieve the next generation of technological goals. Here, while making it smaller and making it better we, unlike our ancestors of the ancient arts, now have both the tools (i.e. TEM, STM, atomic force microscopy (AFM) *et al.*) and the foundational science upon which to 'make it work' in the quest for functional, practical, and effective nanodevices.

12.3 Nanofabrication: creating the building blocks for devices

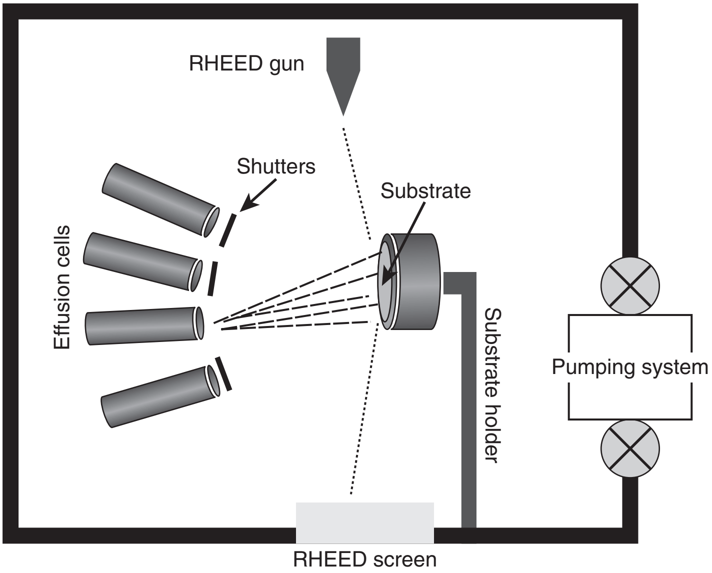
There are as many means for the formation of nanoscale structures as there are types of structures and devices themselves. The basis for any nanodevice is the elemental or compound materials of which it is composed. In Chapter 10 Cheng provides detailed discussion on the subject of nanomaterials. Mention of the same here is simply to provide some context for the fabrication of the devices to be discussed in this chapter. To this end, a few of the primary methods and equipment which have historically contributed to 'thin film growth' and later nanostructures and devices are discussed below.

12.3.1 The source: materials growth

Well before 'nano' became the buzzword for small scale structures, the equipment and methods necessary for the creation of thin films (of a few microns) were being explored. For example, the 1970s saw the use of techniques such as liquid phase epitaxy (LPE) and vapor phase epitaxy (VPE) to ‘grow’ high quality crystalline layers that eventually led to the development of complex devices such as the heterojunction semiconductor laser and compound semiconductor (i.e. GaAs)-based infrared light-emitting diodes (Mishurnyi *et al.*, 2006). Chemical vapor deposition (CVD), whose origins can be traced as far back as the late 1800s, became a staple of silicon thin-film technology, along with the exponential growth of the electronics industry (Allendorf, 1998). These thin-film growth techniques and others evolved through a fairly natural progression from early twentieth century developments in crystal pulling from melts, associated with names such as Czochalski and Bridgman (Scheel and Fukuda, 2003). Today, more refined variations of LPE, VPE, CVD, and other methods are still in use for commercial as well as specialized or custom processing steps in the production of semiconductor thin-film based devices.

However, one materials growth technique has emerged as the ‘workhorse’ of thin-film and nanoscale structures and devices. Known as molecular beam epitaxy (MBE), it has been in use since its inception in 1970 (Cho, 1999). Along with its ‘kissin’ cousin’ metal organic chemical vapor deposition (MOCVD), MBE developed out of the early 1960s in the temperature and vapor pressure control of molecular beams under vacuum and their interaction with solid surfaces. Although there are several variations on the theme (i.e. gas source MBE), in essence traditional MBE involves the use of solid source elemental species of materials which, under carefully controlled conditions of temperature and vacuum, sublime to become the source materials for epitaxial crystal growth. Effusion cells, also known as Knudsen cells or simply ‘furnaces,’ contain these source elements. By adjusting the cell temperatures, a corresponding beam equivalent pressure (BEP) is created, which, in turn, determines the amount of source material deposited on the target substrate. A series of shutters, in the form of non-reactive metal barriers located in front of each cell, can be opened or closed as determined by the MBE operator to select which binary, ternary, or even quaternary combinations of materials become the layers of MBE thin-film growth. Figure 12.2 illustrates the basic equipment configuration necessary for crystal growth via MBE.

The unique and essential characteristic of the MBE technique (and similar methods) is that it can produce both lattice-matched and strained (i.e. pseudomorphic) layers of materials beyond combinations of the same as found in nature. In fact, when parameters such as bandgap energy are taken into account, in addition to the crystalline lattice constants of the materials involved, nanoscale structures such as quantum dots (QD) or quantum wells can be created. The versatility of these and other parameters with respect to the ultimate development of nanodevices is well illustrated in the now commonplace plots of bandgap energy vs lattice constant – an example of which is shown for compound or III-V semiconductors (with reference to columns

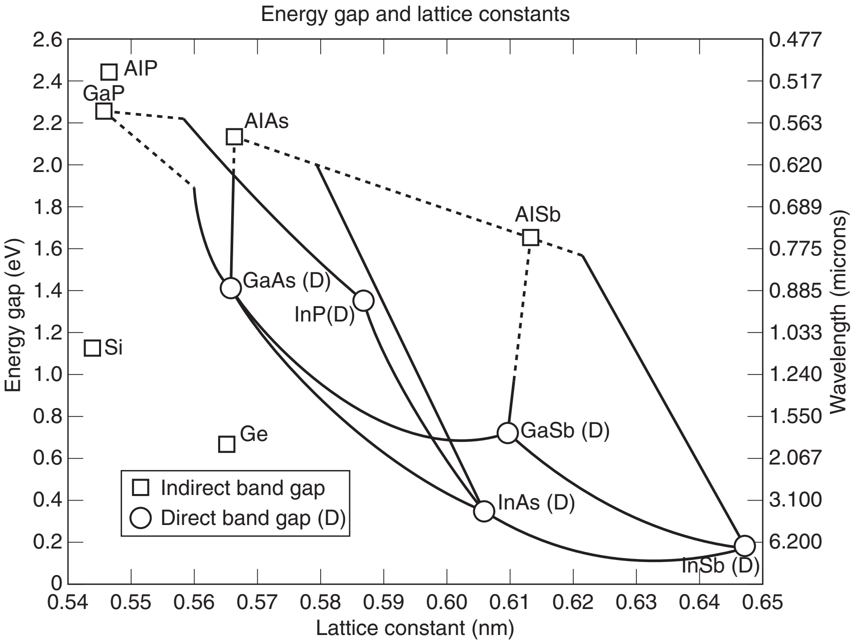


12.2 Schematic of an MBE growth chamber. (Shown with thanks to A. Sarkar for design and drafting).

III and V of the Periodic Table) in Fig. 12.3. From the device perspective, the mole fraction of each binary or higher combination can be selected to create 'bandgap engineered' epitaxial layers, in conjunction with materials of specific emission wavelengths, corresponding to, for example, a desired optoelectronic device functionality.

For this figure, the tie lines between each binary compound represent a ternary combination, as would be formed based on the mole fraction of the respective elements. Thus, $Al_xGa_{1-x}As$, with $x = 0.5$, would lie half way between the AlAs and GaAs tie-line endpoints. Likewise, the plot position of AlAs, practically directly above GaAs, indicates that, in nature, these two binaries are closely lattice-matched, as would be any ternary formed along the line between the two. Quaternary combinations are represented in the areas bounded by the tie lines. For example, $Ga_xIn_{1-x}P_yAs_{1-y}$ is the area bounded by GaP to GaAs and InAs to InP. With commercially and commonly available III-V single crystalline substrates, such as GaAs and InP, Fig. 12.3 can be used to determine the range of lattice-matched and pseudomorphic combinations of epitaxially grown materials available for compound semiconductor device fabrication.

The subject of nanodevices, as it pertains to methods such as MBE, brings to mind structures ranging from the aforementioned QDs to quantum wires and wells, which can be synthesized *in situ* as the natural result of bandgap engineering, lattice matching, and lattice mismatch or strain between the



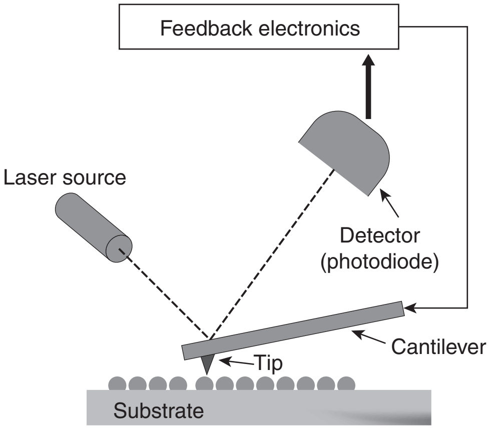
12.3 Traditional plot of bandgap energy and wavelength vs lattice constant for III-V materials. (*Source*: Reproduced with permission from Steve Carson, Executive Relations Director, MIT OpenCourseWare.)

combinations of materials selected. Here again, for the sake of managing a topic of such breadth in the discussion to come, we will later provide an author's choice of select nanoscale device components that have evolved from, not only the 'hard and dry' materials growth techniques such as MBE, but also from other more recently developed vapor–liquid–solid (VLS) combinations of growth methods.

12.3.2 The structure: materials characterization

Since Binning and Roher's realization of the means by which to 'see' atomic structure (Binnig *et al.*, 1982a), a host of scanning probe microscopy (SPM) characterization methods have emerged to make nanodevice development a foreseeable (and currently realizable) process (Bonnell, 2001). Given that many techniques fall under the parent category of SPM, in the spirit of this chapter we briefly mention yet another 'workhorse' of the field, namely AFM. Instrumentation for AFM can be designed to operate in a variety of modes, involving contact, non-contact, and 'tapping' (intermittent contact, oscillation). Other forms of AFM include magnetic force microscopy (MFM), lateral force microscopy (LFM), as well as related techniques such as Kelvin-probe, conductive probe, and capacitance AFM, to name a few. However, these more exotic capabilities essentially evolved from the fundamental ability of AFM to accurately track the topography of a surface regardless of the surface property being measured. A wealth of review articles, textbooks, and even university courses are now available to researchers interested in everything from the origins of AFM to the wide variety of AFM techniques and their applications (Binnig *et al.*, 1986; Hansma *et al.*, 1988; Cappella and Dietler, 1999; Giessibl, 2003; Butt *et al.*, 2005; Muller and Dufrene, 2008; Eaton and West, 2010). In fact, the link www.nanohub.org, as a project funded through the National Science Foundation's Network for Computational Nanotechnology, provides (mostly free of charge) online access to a range instructional tutorials, simulations, and courses in practically all aspects of nanoscience and related fields and technologies. Therefore, in lieu of a detailed historical or technological treatment, we present a brief description of the AFM process to provide some context for materials characterization as a step in the stages of nanodevice development.

Using the interaction of van der Waals forces present between a probe or 'tip' and a sample, AFM instrumentation is designed to translate this nanoscale interaction into real-time 3-D images of the sample surface. Specifically, AFM instrumentation, inclusive of a set of finely tuned computer controlled electric motors (i.e. 'inchworm' electronics), allows placement of the aforementioned tip just a few nanometers from the sample. The attractive–repulsive force interactions between the atoms on the surface and atoms at the end of the AFM tip result in deflections of a cantilever to which the tip is attached, as shown in Fig. 12.4.



12.4 Schematic representation of the main components used to scan a sample using AFM. (Shown with thanks to A. Sarkar for design and drafting.)

Next, a laser focused on the cantilever, along with a mirror–photodetector combination, tracks the physical deflections of the cantilever as the tip is scanned across the surface of the sample. These deflections, as detected by the photodiode, are converted into an electrical signal that is sent to the feedback electronics of the AFM as shown. From this point, an electronic signal sent back to the cantilever-tip positioning electronics (not shown) will adjust parameters such as tip position or frequency of oscillation, depending on the imaging mode.

Finally, the resulting cantilever deflection-to-electronic signal can be converted to an image of the sample's relative atomic structure. The extent of resolution to which an AFM image portrays a given structure depends on everything from the shape of the tip, to tip–sample distance, to AFM mode of operation (i.e. contact vs intermittent or non-contact modes), to control of the natural vibrations in the room in which the instrument is located. Nonetheless, as an SPM technique, AFM, in the hands of a well-trained operator, can serve as a relatively quick and non-destructive means of obtaining nanoscale structural information that can aid in the fabrication of nanodevices.

12.4 Prospects for low dimensional devices

Just as the growth of microelectronics involved breakthroughs and developments in multiple fields, nanoelectronics has been no less varied in terms of its range of constantly emerging new techniques for materials synthesis, characterization, and device fabrication. For the most part, microelectronic devices, particularly early-on the field, can be associated with the realm of 'hard and dry' materials (i.e. semiconductors). On the other hand, the realization of nanoelectronic devices involves both these and other more hybrid methods involving combinations of VLS materials synthesis and processing techniques. With this variety in mind, it is not within the scope of this chapter to discuss the full range of today's nanodevices. However, the examples selected below will, we trust, be reasonably exemplary of the types of nanodevices and their components that have come to fall under the umbrella of 'low dimensional.'

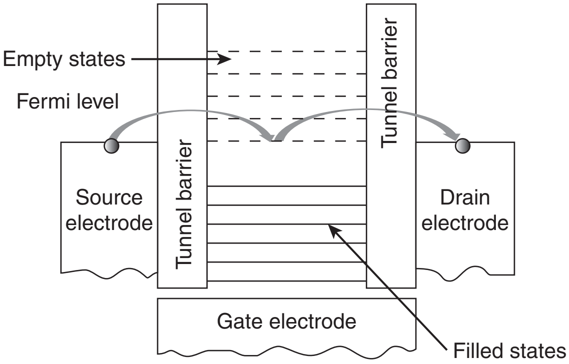
12.4.1 'Dots' to devices: 0th dimension to single-electron transistors

Some of the earliest structures in the nanoparticle regime (i.e. <100 nm in any dimension), known as 'QDs,' were synthesized (grown) via MBE. In general QDs, be they in the solid or colloidal form, are essentially nanoscale structures or particles whose size determines the optoelectronic properties of these entities. As such, electronic confinement within a QD is considered to be of the 'zeroth dimension,' whereas quantum wires and quantum wells scale up 1-D and 2-D electron confinement, respectively. In the early 1990s and going forward, an area of intense research activity developed around the growth of, for example, InAs QDs on GaAs substrates (Brandt *et al.*, 1992; Notzel, 1996; Yoffe, 2001; Joyce and Vvedensky, 2004; Wong *et al.*, 2010). The natural lattice mismatch between these two materials allowed for the creation of nanoparticles in the form of QDs as the InAs deposited, under the previously discussed MBE-controlled conditions, would self-assemble into nanoscale island formations with potentially useful characteristics for electronic and optoelectronic devices. Within the next decade, researchers took the QD, as a device component, well beyond the aforementioned MBE-based work. Thus, variations on the theme, so to speak, of QD studies led to a flurry of research, which, in turn, motivated the development of device prototypes for quantum-dot lasers (Shields, 2007; Bimberg and Pohl, 2011; Lee *et al.*, 2012), sensors and detectors (Posani *et al.*, 2006; Stiff-Roberts, 2009; Li *et al.*, 2011), and solar cells (Nozik, 2002; Tsakalakos, 2008; Pattantyus-Abraham *et al.*, 2010; Jeltsch *et al.*, 2012), to name a few.

However, one nanoparticle-based device that especially generated a great deal of interest was that of the single-electron transistor (SET). Theoretical considerations for single electron charging effects were introduced as early as 1951 (Gorter, 1951), but demonstration of the same did not emerge until the late 1980s to the early 1990s (Barner and Ruggiero, 1987; Kastner, 1992). Given that the 'simple transistor' is the fundamental building block from which the multi-billion dollar electronics industry has grown, it is not surprising that the lion's share of nanoelectronic *device* research has been focused on developing a new generation of transistors and/or transistor-like devices. Thus, built upon the traditional three-terminal source, drain, and gate concept of a field effect transistor (FET), the SET integrates the nanoscale considerations of the *quantum size effect* and *Coulomb blockade* phenomena into its basis of operation. To understand the application of these concepts in an SET a brief explanation is in order.

Single electron transistor (SET): basics of operation

In its essence, quantum size effect describes the accepted theory of quantum mechanics with respect to particle confinement. As the dimensions in which the particle is confined are reduced, a critical limit is reached where the energy level separations normally experienced by this particle are effectively spread even further apart. One may think of 'squeezing' the particle, as in the classic 'particle in a box problem,' but now to the extent where the particle is the size of the box itself. With respect to nanoparticles, this means that particle size now plays a significant role in terms of the electronic characteristics.



12.5 Schematic of single-electron transistor. (Used with thanks to A. Sarkar for design and drafting.)

Coulomb blockade is best examined in the context of the actual SET. To begin, the acronym SET refers to the actual control of electrons, one at a time, between the source and drain of a transistor structure. Figure 12.5 schematically illustrates the configuration of an SET. As shown, it consists of a nanoparticle confined between but separated from direct contact with source and drain by thin insulating (tunneling) barriers. A conductive gate electrode is placed within close proximity to the nanoparticle as well. The nanoparticle contains a total number of N electrons which, in turn, experience discrete energy levels due to the quantum size effect – (two electrons are allowed per energy level, as per Pauli exclusion). The size-controlled electronic configuration is such that the topmost filled state almost aligns with the Fermi level of the source-drain electrodes. In the non-biased state of the SET, there exists a range of empty states available to the electrons of the nanoparticle but not available to any electron outside of the insulating barriers. Theory predicts that if one electron were added to the nanoparticle, such that the total number of electrons becomes $N + 1$, then particle energy would be increased by an amount $q^2/2C$, where C represents the capacitance of the SET. This 'small' change in energy now has a tremendous effect under the aforementioned conditions of quantum confinement, where such a small C translates to a large shift in the available energy states shown in Fig. 12.5. Given this upward shift of available empty states, the opportunity for any quantum mechanical tunneling of electrons from the source electrode is effectively 'blocked' and thus eliminated. However, the single-electron action of the SET comes into play if a voltage of $q/2C$ is applied to the gate electrode. This amount of applied voltage will, in fact, lower the lowest level of the upper empty available states into alignment with the Fermi level of the source electrode. In turn, bias applied across the source and drain electrodes (as per traditional transistor operation) provides the necessary quantum electron conditions for a single electron to 'hop' from the source, to the nanoparticle, and then hop back out through the drain. Note that only one electron at a time can make this journey through the SET given properly applied gate voltage. On the other hand, if two electrons were to make their way onto the nanoparticle at the same time, there would be an increase in capacitance energy which would necessitate increasing the gate voltage even further in order to keep the available states created by the now $N + 2$ electrons in position for the SET process to continue.

SET challenges

As one might imagine, beyond the fabrication challenges of constructing a device whose very operation is built upon the size control of nanoparticles, the fundamental physics of thermal *vs* tunneling energies must be considered. For example, if the capacitance of an SET is too large, then the aforementioned energy shift of $q^2/2C$ will not be large enough to overcome thermal energy. The latter will contribute to the tunneling of electrons between source and drain, thus rendering the voltage controlled gate electrode (i.e., the 'on-off switch' of the transistor) essentially useless. This is particularly an issue for room temperature operation, which is, of course, necessary for any practical transistor applications. However, since the mid 1990s researchers have refined theories and reported on the fabrication of SETs and subsequent transistor performance with varying degrees of success (Takahashi *et al.*, 1995; Matsumoto *et al.*, 1996; Feldheim and Keating, 1998; Goldhaber-Gordon *et al.*, 1998; Knobel and Cleland, 2003; LaHaye *et al.*, 2004; Shekhter *et al.*, 2007; Cheam *et al.*, 2009; Khondaker *et al.*, 2010; Maeda *et al.*, 2012). Thus, the SET serves as a prime example of nanodevice operation via nanoparticle-based phenomena.

12.4.2 Nano(tubes)wires: from 1-D to circuits

As mentioned in Section 12.1, the breakthrough discovery of C_{60} led to the likewise milestone observation of CNTs. First identified via electron microscopy (Iijima, 1991), CNTs have captivated practically every realm of the scientific community with potential applications ranging from drug delivery to cancer cells to materials reinforcement as per their tensile 'stronger than steel' properties (Ji *et al.*, 2010; Gomez-Gualdron *et al.*, 2011; Jiang *et al.*, 2011). However, as far as their use in electronic devices is concerned, CNTs provide a means by which to address the Moore's Law dilemma faced by traditional semiconductor technologies in terms of increasing device density while decreasing device feature size. Even the 'wiring' as per interconnects between devices becomes a substantial factor of technological concern. Eventually one must face the challenges of physical phenomena (i.e. quantum size effect) as well as economic considerations (i.e. cost of processing and fabrication technologies) which, in traditional semiconductor materials, may far outweigh the advantages of shrinking device dimensions. Thus new nanoscale-based components for fabricating CNTs are being investigated for feasibility in next-generation nanoelectronic devices (Avouris *et al.*, 2007; Naeemi and Meindl, 2009; Lee *et al.*, 2010; Zhang *et al.*, 2012b).

Carbon nanotube (CNT): device integration

One early example of CNT nanowire technology was explored by Postma, *et al.* who essentially fabricated an SET but with the nanoparticle replaced by a CNT (Postma *et al.*, 2001). (Given its structure and effective function, we will refer to this 'tube' as a CNT nanowire in the description to follow.) An AFM was used to strategically 'kink' the nanowire in such a way as to create the tunneling barriers required for SET-type operation. The conductance measured for this structure could be modulated via an applied gate voltage. Likewise, the device itself exhibited a Coulomb gap of 120 meV, meaning that some measure of on–off threshold voltage control was achieved.

Another first step in the use of CNTs for electronic devices was made by Bachtold *et al.*, who used CNT nanowire based transistors, but now 'integrated,' in a proof-of-concept for the development of circuits (Bachtold *et al.*, 2001). In this work, a circuit consisting of up to three transistors was fabricated, such that each device was controlled by its own gate electrode. Unlike the particle-based SET, these devices were fabricated in the form of the traditional FET whose channel width was equal to the length of a single CNT between Au source and drain contacts. The voltage applied to each independent gate electrode could be used to shift the Fermi level of each device through its respective valence band, bandgap, or conduction band, so that any given transistor in this circuit would exhibit p-type (hole) or n-type (electron) conduction. Ultimately, one can imagine a logic operation where, for example, two CNT FETs are configured to act as a three-terminal device (two inputs, one output), as in a NOR gate. Given the independence in gate electrode control, as described previously, and using the applied voltages to act as the equivalent of a 'logic 1' or 'logic 0,' this CNT-based structure effectively achieved the same 'universal gate' performance as with microelectronic circuits where specifically configured sets of NOR gates can be used to replicate the behavior of any other type of digital logic gate.

CNT: challenges

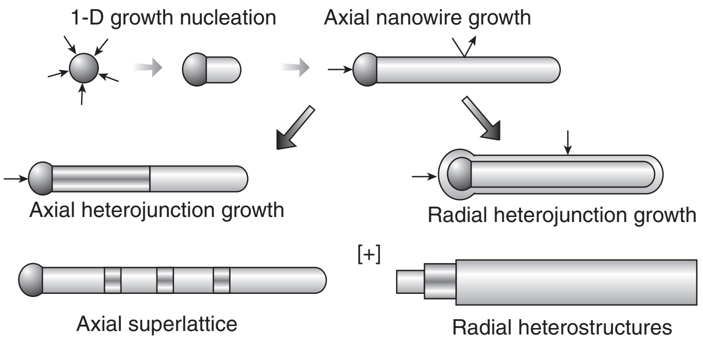
Going forward, researchers have found that some of the major challenges for CNT-based devices were not so much with the CNTs themselves but with the sources, drains, and interconnects used for transistor-to-circuit level development. In the above examples, even these groundbreaking CNT transistors and circuits required the use of traditional top-down technology as the overall structures were juxtaposed within the same scale as that of their microelectronic counterparts. Nonetheless, the past decade has seen tremendous progress in the integration of CNTs into circuit architectures (Chau *et al.*, 2005; Graham *et al.*, 2005; Chen *et al.*, 2006; Kang *et al.*, 2007; Cao *et al.*, 2008; Yan *et al.*, 2011; Ding *et al.*, 2012). In fact, recent publications and reviews on CNT-based devices and circuits describe the ongoing work in the application of CNTs in a variety of electronic device areas, from thin-film transistor circuits (Pribat and Bondavalli, 2012) to models for next-generation very large-scale integration (VLSI) designs (Zhang *et al.*, 2012a). As a result, CNTs have set the stage for future nanodevices where nanowire technology is being explored.

12.5 Beyond the bottom-up: hybrid nanoelectronics

With recognition of the seminal breakthroughs that have emerged from CNTs, as discussed above, a likewise exciting range of devices and circuits has been under study beyond this particular form of single element electronics. Specifically, investigations into the physics, chemistry, and electronic functionality of nanowires, nanobelts, and related nanostructures for devices have increased exponentially since the late 1990s (Hu *et al.*, 1999; Pan *et al.*, 2001; Wang, 2004; Lieber and Wang, 2007; Fang *et al.*, 2011). Moreover, as mentioned in Section 12.3.1, in addition to traditional epitaxial growth techniques, the principles of vapor deposition have come full circle as combinations of VLS growth techniques have been developed in the search for 1-D nanowires and other low-dimensional structures (Trentler *et al.*, 1995; Yang *et al.*, 2002; Law *et al.*, 2004; Lieber and Wang, 2007; Tsivion *et al.*, 2011). For example, in VLS growth, a nanodroplet sized catalyst is used as the basis for everything from preferentially directing the reactant species that ultimately forms the wire to determining the nanoscale diameter of the 1-D structure itself (Lieber and Wang, 2007). However, moving from the 'interesting chemistry' of the aforementioned to electronic devices and circuits opens up a range of device fabrication techniques – either top-down, bottom-up, or a hybrid of the two. A few examples of these next-generation low-dimension device components have been selected for discussion below.

12.5.1 Heterostructure nanowires: the building blocks

Once again the versatility of 1-D nanoelectronics provides a particularly promising range of device technologies in the form of *heterostructure*



12.6 Schematic of nanowire growth for both axial (left) and radial (right) heterostructures. (*Source*: Reprinted from *Nano Today*, 3, Hayden O, Agarwal R and Lu W, Semiconductor nanowire devices, 12-22, 2008 with permission from Elsevier. Also reproduced with the permission of Professor Wei Lu (Hayden *et al.*, 2008).)

nanowires (HNW). To first order, one can divide HNWs into two types: axial (superlattice) and radial (core–shell) formations. Several examples of both are detailed in the literature, but briefly, both structures are synthesized using some form of the aforementioned VLS method of materials growth. The axial HNW structure involves vapor adsorption and/or decomposition at the surface of a nanocluster catalyst site. Selective changes in the reactants needed for growth of the nanowire (NW) will create a controlled heterostructure formation as per Fig. 12.6.

Alternatively, if the surface of the NW is essentially used as the reactant site, then the adsorption/decomposition process will create a shell of new material concentrically enclosing the original NW. Repetition of this 'surface reaction' can produce multishell radii of NWs surrounding the originally grown core NW, also shown in Fig. 12.6. With even these two primary forms of nanowire synthesis, the phrase 'the possibilities are (almost) endless' comes to mind (Lauhon *et al.*, 2002).

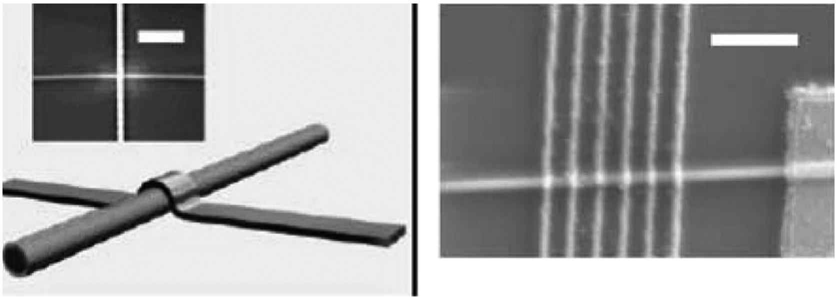
One such device possibility, for the axial or superlattice HNW, is that of the SET (Section 12.4.1). For example, in 2003 InAs/InP HNWs, seeded by metal (Au) catalyst particles and with InP double barriers, showed excellent SET characteristics at low temperature (4.2 K) (Thelander *et al.*, 2003). However, as is often the case with new device technologies, contacts to those devices (i.e. source/drain) may be as challenging, if not more so, than the fabrication of the core device structure itself. Thus, taking device feasibility one step further, the investigators also structured their HNW SET such that the transistor action of the structure would be effectively independent of contact design. In 2008, the first demonstration of an SET operated in rf mode was reported (Nilsson *et al.*, 2008). In this work, the investigators epitaxially grew InAs/InP HNWs from Au aerosol nanoparticles and reported transistor operation at 1.5 K with typical $1/f$ behavior at low frequencies. Since then, axial nanowires have been investigated for their feasible use in devices such as tunnel diodes (Wallentin *et al.*, 2010), in FETs (Tomioka and Fukui, 2011), as photovoltaic components (Tian and Lieber, 2011), and as localized biosensor probes (Jiang *et al.*, 2012).

For radial or core-shell HNWs, the range of nanodevice development may be even greater. For example, investigators from the Lieber Group have used undoped Ge/Si core/shell HNWs between a set of superconducting leads to create a device in which proximity-induced superconductivity could be observed (Xiang *et al.*, 2006b). In this device, critical supercurrent was tuned from 0 to over 100 nA via modulation of carrier density in the Ge/Si HNW via a top gate electrode. This group has also used Ge/Si HNWs to fabricate quantum devices exhibiting near ideal performance due to charge states in coupled QDs formed inside of the NWs (Hu *et al.*, 2007). Recent studies of Si/Ge HNWs have investigated properties related to tunneling (Nah *et al.*, 2010; Fung *et al.*, 2011; Hao *et al.*, 2011;) and spin relaxation (Hu *et al.*, 2012). Radial HNWs synthesized in the III-V material system have been no less interesting. Comprehensive studies of III-V based NWs can be found in articles such as the work by Joyce *et al.* (2011). More device specific III-V NW investigations include studies of core multishell NWs for light-emitting diodes (Tomioka *et al.*, 2010) and as photovoltaic components (Borgstrom *et al.*, 2011).

The engagement of NWs (heterostructure or single element) as the primary components in circuit architectures will be needed to take the field from the realm of 'proof-of-concept' to the revolutionary next step 'beyond the transistor' in the development of nanoelectronic technology. As was mentioned at the start of this chapter, breakthroughs toward the successful realization of applications ranging from nanocomputing, to consumer electronics, to biological *in vivo* detection, will require advances beyond synthesis and the self-assembly of these and related nanoscale electronic structures. Thus, the section to follow provides a look at a thus-far promising approach toward achievement of this ultimate goal of nanoelectronic circuit functionality.

12.5.2 Nanowire arrays – the circuit regime

Just as nanoelectronic device development can be categorized in terms of low-dimensional functionality, nanoscale circuit realization is following a similar path. For the sake of illustrating this progression, we continue the discussion of this chapter in terms of NWs but now from the perspective of circuit architecture. An interesting example of the latter can be found in a work of Lieber and co-workers in the discussion to follow.



12.7 Schematic of nanowire crossbar configuration (left), SEM inset crossed NW scale bar is 1 um; SEM of nanowire crossbar (1 × 6) array (right), scale bar is 500 nm. (*Source*: Reprinted with permission from Dong Y J, Yu G H, McAlpine M C, Lu W and Lieber C M 2008 Si/a-Si core/shell nanowires as nonvolatile crossbar switches *Nano Letters* 8 386–389. Copyright 2008 American Chemical Society.)

Key necessary elements for circuit realization are available via the core–shell configuration of NWs (Dong *et al.*, 2008). To begin, as described in Section 12.5.1, the selection of core–shell materials determines the electrical characteristics of the given NWs. Using the synthesis of NWs as 'building blocks,' a crossbar configuration can be used as a scalable architecture from which to build the higher functions of logic and memory in nanoelectronics. Figure 12.7 illustrates the use of the crossbar configuration for the demonstration of a hysteric resistance switch, built using a core–shell Si NW, for the realization of a 1-D memory array. Details concerning the construction of this circuit element can be found in the published work of Yu and Lieber and references therein (Yu and Lieber, 2010). However, to the point of nanoscale circuit development, it is worth pointing out that this switch exhibited the significant features of bistable switching, reproducible switching threshold voltages, and ON state current rectification – all under room temperature operation.

Taking the above NW crossbar device to the next step into the circuit regime, Lieber and co-workers also fabricated (1 × 6) small scale 1-D arrays (Fig. 12.7-right). This nano *circuit* structure exhibited write/erase and read-out capability, without crosstalk, from the six switches. Thereafter, to push the limits of the crossbar NW switch to higher dimensional functionality, these same investigators fabricated highly integrated 2-D arrays (Yu and Lieber, 2010). The subsequently fabricated nanocircuit required the use of both top-down and bottom-up technologies in the form of high-resolution electron-beam lithography and the well-controlled synthesis of two sets of parallel high-density NW arrays, respectively. Working demonstration of the same was accomplished using a (6 × 6) crossbar nanowire array which, as with the 1-D array, successfully exhibited read/write and memory capabilities.

Finally, true nanoelectronic circuitry will require the realization of 3-D integration based on nanodevices. Toward this goal, significant progress has been made as well (Stan *et al.*, 2003; Zhang and Chou, 2003; McAlpine *et al.*, 2007; Jo *et al.*, 2009; Shiratori *et al.*, 2011; Wierer *et al.*, 2012). Following through from the aforementioned 1-D and 2-D work, a particularly interesting and successful foray into 3-D realization is found based on the use of nanoimprinting (Bao *et al.*, 2002; Guo, 2007). Using this technique, in conjunction with traditional device fabrication, Lieber and co-workers were able to achieve vertical stacking of 3-D devices having up to ten addressable (vertical) layers (Javey *et al.*, 2007). The circuit itself was based on an inverter/floating-gate memory structure using Ge/Si core/shell NWs (Xiang *et al.*, 2006a; Yu and Lieber, 2010).

Ultimately, the journey from 'building blocks,' be they particles, wires, or devices, to actual circuit integration is a non-trivial one, to say the least. Challenges remain in terms of the synthesis, ordering, and assembly of dense NW configurations. Yet, presumably, as these present-day obstacles are overcome, the use of nanodevices in circuitry holds the key to a new era of high performance, faster speeds, and a wider variety of applications for the next generation of the electronics industry.

12.6 Conclusion and future trends

As one might expect by virtue of this chapter being written and included in this book, the outlook for nanodevices is certainly promising. In fact, as has been pointed out several times, it would take more than a chapter to even begin to cover the full range of this subject. Thus, this author has worked to provide at least a template for further study, on the part of the reader, by selecting examples of some of the major milestones in the continuing successful development of nanodevices and the structures of which they are composed.

That being said, challenges do remain in terms of the comprehensive development of nanoscale device components within practical 3-D circuitry. Interconnects, device integration, and basic reliability are ongoing areas of intensive research. Among the many 'proof-of-concept' structures, nanodevices, albeit on a positive trajectory, are still closer to the fundamental scientific development end of the spectrum as opposed to more immediate commercial application. Nevertheless, in light of the many breakthroughs, as have been presented, even the most conservative outlook towards nanodevice development would encourage ongoing research and development down this path toward the 'beyond the transistor' generation of electronics.

12.7 References

Ahn J H and Je J H (2012), Stretchable electronics: Materials, architectures and integrations, *J. Phys. D-Appl. Phys.,* **45**, 14.

Allendorf M (1998), From Bunsen to VLSI: 150 years of growth in chemical vapor deposition technology, *Interface,* **7**, **1**–3.

Avouris P, Chen Z H and Perebeinos V (2007), Carbon-based electronics, *Nat. Nanotechnol.,* **2,** 605–615.

Bachtold A, Hadley P, Nakanishi T and Dekker C (2001), Logic circuits with carbon nanotube transistors, *Science,* **294**, 1317–1320.

Bao L R, Cheng X, Huang X D, Guo L J, Pang S W and Yee A F (2002), Nanoimprinting over topography and multilayer three-dimensional printing, *J. Vac. Sci. Technol. B,* **20,** 2881–2886.

Barner J B and Ruggiero S T (1987), Observation of the incremental charging of Ag-particles by single electrons, *Phys. Rev. Lett.,* **59**, 807–810.

Bimberg D and Pohl U W (2011), Quantum dots: Promises and accomplishments, *Mater. Today,* **14**, 388–397.

Binnig G, Quate C F and Gerber C (1986), Atomic force microscope, *Phys. Rev. Lett.,* **56**, 930–933.

Binnig G, Rohrer H, Gerber C and Weibel E (1982a) Surface studies by scanning tunneling microscopy, *Phys. Rev. Lett.,* **49**, 57–61.

Binnig G, Rohrer H, Gerber C and Weibel E (1982b) Tunneling through a controllable vacuum gap, *Appl. Phys. Lett.,* **40**, 178–180.

Binns C (2010), *Introduction to Nanoscience and Nanotechnology*: (John Wiley & Sons, Inc.), Hoboken, New Jersey.

Bonnell D ed (2001), *Scanning Probe Microscopy and Spectroscopy*: (Wiley-VCH), Weinheim, Germany.

Borgstrom M T, Wallentin J, Heurlin M, Falt S, Wickert P, Leene J, Magnusson M H, Deppert K and Samuelson L (2011), Nanowires with promise for photovoltaics, *IEEE J. Sel. Top.Quantum Electron.,* **17**, 1050–1061.

Brandt O, Tapfer L, Ploog K, Bierwolf R, Phillipp F and Hohenstein M (1992), Direct synthesis of InAs quantum dots in single-crystalline GaAs matrix by molecular-beam epitaxy, *Surf. Sci.,* **267**, 204–208.

Butt H J, Cappella B and Kappl M (2005), Force measurements with the atomic force microscope: Technique, interpretation and applications, *Surf. Sci. Rep.,* **59**, 1–152.

Cao Q, Kim H S, Pimparkar N, Kulkarni J P, Wang C J, Shim M, Roy K, Alam M A and Rogers J A (2008), Medium-scale carbon nanotube thin-film integrated circuits on flexible plastic substrates, *Nature,* **454**, 495–500.

Cappella B and Dietler G (1999), Force-distance curves by atomic force microscopy, *Surf. Sci. Rep.,* **34**, 1–3, 5–104.

Chau R, Datta S, Doczy M, Doyle B, Jin J, Kavalieros J, Majumdar A, Metz M and Radosavljevic M (2005), Benchmarking nanotechnology for high-performance and low-power logic transistor applications, *IEEE Trans. Nanotechnol.,* **4**, 153–158.

Cheam D D, Karre P S K, Palard M and Bergstrom P L (2009), Step and flash imprint lithography for quantum dots based room temperature single electron transistor fabrication, *Microelectron. Eng.,* **86**, 646–649.

Chen Z H, Appenzeller J, Lin Y M, Sippel-Oakley J, Rinzler A G, Tang J Y, Wind S J, Solomon P M and Avouris P (2006), An integrated logic circuit assembled on a single carbon nanotube, *Science,* **311**, 1735.

Cho A Y (1999), How molecular beam epitaxy (MBE) began and its projection into the future, *J. Cryst. Growth,* **201**, 1–7.

Ding L, Zhang Z Y, Liang S B, Pei T, Wang S, Li Y, Zhou W W, Liu J and Peng L M (2012), CMOS-based carbon nanotube pass-transistor logic integrated circuits, *Nat. Commun.,* **3**, 7.

Dong Y J, Yu G H, McAlpine M C, Lu W and Lieber C M (2008), Si/a-Si core/shell nanowires as nonvolatile crossbar switches, *Nano Lett.,* **8**, 386–391.

Drexler K E (1992), *Nanosystems: Molecular Machinery, Manufacturing, and Computation*: (Wiley Interscience), Hoboken, New Jersey.

Eaton P and West P (2010), *Atomic Force Microscopy*: (Oxford University Press, USA).

Fang X S, Zhai T Y, Gautam U K, Li L, Wu L M, Yoshio B and Golberg D (2011), ZnS nanostructures: From synthesis to applications, *Prog. Mater. Sci.,* **56**, 175–287.

Feldheim D L and Keating C D (1998), Self-assembly of single electron transistors and related devices, *Chem. Soc. Rev.,* **27**, 1–12.

Feyman R P (1960), There's plenty of room at the bottom, *Caltech Eng. Sci.,* **23**, 22–36.

Fung W Y, Chen L and Lu W (2011), Esaki tunnel diodes based on vertical Si-Ge nanowire heterojunctions, *Appl. Phy. Lett.,* **99**, 3.

Giessibl F J (2003), Advances in atomic force microscopy, *Rev. Mod. Phys.,* **75**, 949–983.

Goldhaber-Gordon D, Shtrikman H, Mahalu D, Abusch-Magder D, Meirav U and Kastner M A (1998), Kondo effect in a single-electron transistor, *Nature,* **391**, 156–159.

Gomez-Gualdron D A, Burgos J C, Yu J M and Balbuena P B (2011), *Progress in Molecular Biology and Translational Science: Nanoparticles in Translational Science and Medicine,* ed A Villaverde (San Diego: Elsevier Academic Press Inc) pp 175–245.

Gorter C J (1951), A possible explanation of the increase of the electrical resistance of thin metal films at low temperatures and small field strengths, *Physica,* **17**, 777–778.

Graham A P, Duesberg G S, Hoenlein W, Kreupl F, Liebau M, Martin R, Rajasekharan B, Pamler W, Seidel R, Steinhoegl W and Unger E (2005), How do carbon nanotubes fit into the semiconductor roadmap? *Appl. Phys. A-Mater. Sci. Process.,* **80**, 1141–1151.

Guo L J (2007), Nanoimprint lithography: Methods and material requirements, *Adv. Mater.,* **19**, 495–513.

Hansma P K, Elings V B, Marti O and Bracker C E (1988), Scanning tunneling microscopy and atomic force microscopy – application to biology and technology, *Science,* **242**, 209–216.

Hao X J, Li H O, Tu T, Zhou C, Cao G, Guo G C, Guo G P, Fung W Y, Ji Z Q and Lu W (2011), Andreev tunneling enhanced by Coulomb oscillations in superconductor-semiconductor hybrid Ge/Si nanowire devices, *Phys. Rev. B,* **84**, 6.

Hayden O, Agarwal R and Lu W (2008), Semiconductor nanowire devices, *Nano Today,* **3**, 12–22.

http://www.microscopy.ethz.ch/history.htm.

Hu J T, Odom T W and Lieber C M (1999), Chemistry and physics in one dimension: Synthesis and properties of nanowires and nanotubes, *Acc. Chem. Res.,* **32**, 435–445.

Hu Y J, Churchill H O H, Reilly D J, Xiang J, Lieber C M and Marcus C M (2007), A Ge/Si heterostructure nanowire-based double quantum dot with integrated charge sensor, *Nat. Nanotechnol.,* **2**, 622–625.

Hu Y J, Kuemmeth F, Lieber C M and Marcus C M (2012), Hole spin relaxation in Ge-Si core-shell nanowire qubits, *Nat. Nanotechnol.,* **7**, 47–50.

Iijima S (1991), Helical microtubules of graphitic carbon, *Nature,* **354**, 56–58.

Javey A, Nam S, Friedman R S, Yan H and Lieber C M (2007), Layer-by-layer assembly of nanowires for three-dimensional, multifunctional electronics, *Nano Lett.,* **7**, 773–777.

Jeltsch K F, Schadel M, Bonekamp J B, Niyamakom P, Rauscher F, Lademann H W A, Dumsch I, Allard S, Scherf U and Meerholz K (2012), Efficiency enhanced hybrid solar cells using a blend of quantum dots and nanorods, *Adv. Funct. Mater.,* **22**, 397–404.

Ji S R, Liu C, Zhang B, Yang F, Xu J, Long J A, Jin C, Fu D L, Ni Q X and Yu X J (2010), Carbon nanotubes in cancer diagnosis and therapy, *Biochim. Biophys. Acta-Rev. Cancer,* **1806**, 29–35.

Jiang K L, Wang J P, Li Q Q, Liu L A, Liu C H and Fan S S (2011), Superaligned carbon nanotube arrays, films, and yarns: A road to applications, *Adv. Mater.,* **23**, 1154–1161.

Jiang Z, Qing Q, Xie P, Gao R X and Lieber C M (2012), Kinked p-n junction nanowire probes for high spatial resolution sensing and intracellular recording, *Nano Lett.,* **12**, 1711–1716.

Jo S H, Kim K H and Lu W (2009), High-density crossbar arrays based on a Si memristive system, *Nano Lett.,* **9**, 870–874.

Joyce B A and Vvedensky D D (2004), Self-organized growth on GaAs surfaces, *Mater. Sci. Eng. R-Rep.,* **46**, 127–176.

Joyce H J, Gao Q, Tan H H, Jagadish C, Kim Y, Zou J, Smith L M, Jackson H E, Yarrison-Rice J M, Parkinson P and Johnston M B (2011), III-V semiconductor nanowires for optoelectronic device applications, *Prog. Quantum Electron.,* **35**, 23–75.

Kang S J, Kocabas C, Ozel T, Shim M, Pimparkar N, Alam M A, Rotkin S V and Rogers J A (2007), High-performance electronics using dense, perfectly aligned arrays of single-walled carbon nanotubes, *Nat. Nanotechnol.,* **2**, 230–236.

Kastner M A (1992), The single electron transistor, *Rev. Mod. Phys.,* **64**, 849–858.

Khondaker S I, Luo K and Yao Z (2010), The fabrication of single-electron transistors using dielectrophoretic trapping of individual gold nanoparticles, *Nanotechnology,* **21**, 4.

Knobel R G and Cleland A N (2003), Nanometre-scale displacement sensing using a single electron transistor, *Nature,* **424**, 291–293.

LaHaye M D, Buu O, Camarota B and Schwab K C (2004), Approaching the quantum limit of a nanomechanical resonator, *Science,* **304**, 74–77.

Lauhon L J, Gudiksen M S, Wang C L and Lieber C M (2002), Epitaxial core-shell and core-multishell nanowire heterostructures, *Nature,* **420**, 57–61.

Law M, Goldberger J and Yang P D (2004), Semiconductor nanowires and nanotubes, *Ann. Rev. Mater. Res.,* **34**, 83–122.

Lee J, Oszwaldowski R, Gothgen C and Zutic I (2012), Mapping between quantum dot and quantum well lasers: From conventional to spin lasers, *Phys. Rev. B,* **85**, 13.

Lee S W, Yabuuchi N, Gallant B M, Chen S, Kim B S, Hammond P T and Shao-Horn Y (2010), High-power lithium batteries from functionalized carbon-nanotube electrodes, *Nat. Nanotechnol.,* **5**, 531–537.

Li L L, Liu K P, Yang G H, Wang C M, Zhang J R and Zhu J J (2011), Fabrication of graphene-quantum dots composites for sensitive electrogenerated chemiluminescence immunosensing, *Adv. Funct. Mater.,* **21**, 869–878.

Lieber C M and Wang Z L (2007), Functional nanowires, *MRS Bull.,* **32**, 99–108.

Maeda K, Okabayashi N, Kano S, Takeshita S, Tanaka D, Sakamoto M, Teranishi T and Majima Y (2012), Logic operations of chemically assembled single-electron transistor, *Acs Nano,* **6**, 2798–2803.

Matsumoto K, Ishii M, Segawa K, Oka Y, Vartanian B J and Harris J S (1996), Room temperature operation of a single electron transistor made by the scanning tunneling microscope nanooxidation process for the TiOx/Ti system, *Appl. Phys. Lett.,* **68**, 34–36.

McAlpine M C, Ahmad H, Wang D W and Heath J R (2007), Highly ordered nanowire arrays on plastic substrates for ultrasensitive flexible chemical sensors, *Nat. Mater.,* **6**, 379–384.

Mishurnyi V A, de Anda F, Elyukhin V A and Hernandez I C (2006), Growth of quantum-well heterostructures by liquid phase epitaxy, *Crit. Rev. Solid State Mat. Sci.,* **31**, 1–13.

Moore G (2005), Excerpts from A Conversation with Gordon Moore: Moore's Law. (Intel Corporation).

Moore G E (1998), Cramming more components onto integrated circuits (Reprinted from Electronics, pp. 114–117, 19 April 1965), *Proc. IEEE,* **86**, 82–85.

Muller D J and Dufrene Y F (2008), Atomic force microscopy as a multifunctional molecular toolbox in nanobiotechnology, *Nat. Nanotechnol.,* **3**, 261–269.

Naeemi A and Meindl J D (2009), Carbon nanotube interconnects, *Ann. Rev. Mater. Res.,* (Palo Alto: Annual Reviews), **39**, 255–275.

Nah J, Liu E S, Varahramyan K M and Tutuc E (2010), Ge-Si(x)Ge(1-x) Core-shell nanowire tunneling field-effect transistors, *IEEE Trans. Electron Devices,* **57**, 1883–1888.

Nilsson H A, Duty T, Abay S, Wilson C, Wagner J B, Thelander C, Delsing P and Samuelson L (2008), A radio frequency single-electron transistor based on an InAs/InPheterostructure nanowire, *Nano Lett.,* **8** 872–875.

Notzel R (1996), Self-organized growth of quantum-dot structures, *Semicond. Sci. Technol,.* **11**, 1365–1379.

Nozik A J (2002), Quantum dot solar cells, *Physica E,* **14**, 115–120.

Pan Z W, Dai Z R and Wang Z L (2001), Nanobelts of semiconducting oxides, *Science,* **291**, 1947–1949.

Pattantyus-Abraham A G, Kramer I J, Barkhouse A R, Wang X, Konstantatos G, Debnath R, Levina L, Raabe I, Nazeeruddin M K, Graetzel M and Sargent E H (2010), Depleted-heterojunction colloidal quantum dot solar cells, *Acs Nano,* **4**, 3374–3380.

Posani K T, Tripathi V, Annamalai S, Weisse-Bernstein N R, Krishna S, Perahia R, Crisafulli O and Painter O J (2006), Nanoscale quantum dot infrared sensors with photonic crystal cavity, *Appl. Phys. Lett.,* **88**, 3.

Postma H W C, Teepen T, Yao Z, Grifoni M and Dekker C (2001), Carbon nanotube single-electron transistors at room temperature, *Science,* **293**, 76–79.

Pribat D and Bondavalli P (2012), Thin-film transistors and circuits based on carbon nanotubes, *J. Disp. Technol.,* **8**, 54–60.

Ratner M and Ratner D (2003), *Nanotechnology: A Gentle Introduction to the Next Big Idea*: (Prentice Hall Professional), Upper Saddle River, New Jersey.

Renn O and Roco M C (2006) Nanotechnology and the need for risk governance, *J. Nanopart. Res.,* 8, 153–91.

Scheel H J and Fukuda T (2003), *Crystal Growth Technology*: (John Wiley & Sons, Ltd.), Chichester, West Sussex PO19 8SQ, England.

Shekhter R I, Gorelik L Y, Jonson M, Galperin Y M and Vinokur V M (2007), Nanomechanical shuttle transfer of electrons, *J. Comput. Theor. Nanosci.,* **4**, 860–895.

Shields A J (2007), Semiconductor quantum light sources, *Nat. Photonics,* **1**, 215–223.

Shiratori Y, Miura K and Kasai S (2011), Programmable nano-switch array using SiN/GaAs interface traps on a GaAs nanowire network for reconfigurable BDD logic circuits, *Microelectron. Eng.,* **88**, 2755–2758.

Stan M R, Franzon P D, Goldstein S C, Lach J C and Ziegler M M (2003), Molecular electronics: From devices and interconnect to circuits and architecture, *Proc. IEEE,* **91**, 1940–1957.

Stiff-Roberts A D (2009), Quantum-dot infrared photodetectors: A review, *J. Nanophoton.,* **3**, 031607.

Takahashi Y, Nagase M, Namatsu H, Kurihara K, Iwdate K, Nakajima K, Horiguchi S, Murase K and Tabe M (1995), Fabrication technique for Si single-electron transistor operation at room-temperature, *Electron. Lett.,* **31**, 136–137.

Thelander C, Martensson T, Bjork M T, Ohlsson B J, Larsson M W, Wallenberg L R and Samuelson L (2003), Single-electron transistors in heterostructure nanowires, *Appl. Phys. Lett.,* **83**, 2052–2054.

Tian B Z and Lieber C M (2011), Design, synthesis, and characterization of novel nanowire structures for photovoltaics and intracellular probes, *Pure Appl. Chem.,* **83**, 2153–2169.

Tomanek D, Jorio A, Dresselhaus M S and Dresselhaus G (2008), Introduction to the important and exciting aspects of carbon-nanotube science and technology, *Topics Appl. Phys.: Carbon Nanotubes,* **111**, 1–12.

Tomioka K and Fukui T (2011), Tunnel field-effect transistor using InAs nanowire/ Si heterojunction, *Appl. Phys. Lett.,* **98**, 3.

Tomioka K, Motohisa J, Hara S, Hiruma K and Fukui T (2010), GaAs/AlGaAs core multishell nanowire-based light-emitting diodes on Si, *Nano Lett.,* **10**, 1639–1644.

Trentler T J, Hickman K M, Goel S C, Viano A M, Gibbons P C and Buhro W E (1995), Solution-liquid-solid growth of crystalline III-V semiconductors-An analogy to vapor-liquid-solid growth, *Science,* **270**, 1791–1794.

Tsakalakos L (2008), Nanostructures for photovoltaics, *Mater. Sci. Eng. R-Rep.,* **62**, 175–189.

Tsivion D, Schvartzman M, Popovitz-Biro R, von Huth P and Joselevich E (2011), Guided growth of millimeter-long horizontal nanowires with controlled orientations, *Science,* **333**, 1003–1007.

Wallentin J, Persson J M, Wagner J B, Samuelson L, Deppert K and Borgstrom M T (2010), High-performance single nanowire tunnel diodes, *Nano Lett.,* **10**, 974–979.

Walter P, Martinetto P, Tsoucaris G, Breniaux R, Lefebvre M A, Richard G, Talabot J and Dooryhee E (1999), Making make-up in ancient Egypt, *Nature,* **397**, 483–484.

Walter P, Welcomme E, Hallegot P, Zaluzec N J, Deeb C, Castaing J, Veyssiere P, Breniaux R, Leveque J-L and Tsoucaris G (2006), Early use of PbS nanotechnology for an ancient hair dyeing formula, *Nano Lett.,* **6**, 2215–2219.

Wang Z L (2004), Zinc oxide nanostructures: Growth, properties and applications, *J. Phys.-Condes. Matter,* **16**, R829–R858.

Wierer J J, Jr., Li Q, Koleske D D, Lee S R and Wang G T (2012), III-nitride core-shell nanowire arrayed solar cells, *Nanotechnology,* **23**, 194007.

Wille K and Loh K J (2010), Nanoengineering ultra-high-performance concrete with multiwalled carbon nanotubes, *Transp. Res. Record,* 119–126.

Wong P-S, Liang B and Huffaker D L (2010), InAs quantum dots on nanopatterned GaAs (001) surface: The growth, optical properties, and device implementation, *J. Nanosci. Nanotechnol.,* **10**, 1537–1550.

Xiang J, Lu W, Hu Y J, Wu Y, Yan H and Lieber C M (2006a) Ge/Si nanowire heterostructures as high-performance field-effect transistors, *Nature,* **441**, 489–493.

Xiang J, Vidan A, Tinkham M, Westervelt R M and Lieber C M (2006b) Ge/Si nanowire mesoscopic Josephson junctions, *Nat. Nanotechnol.,* **1**, 208–213.

Yan H, Choe H S, Nam S W, Hu Y J, Das S, Klemic J F, Ellenbogen J C and Lieber C M (2011), Programmable nanowire circuits for nanoprocessors, *Nature,* **470**, 240–244.

Yang P D, Yan H Q, Mao S, Russo R, Johnson J, Saykally R, Morris N, Pham J, He R R and Choi H J (2002), Controlled growth of ZnO nanowires and their optical properties, *Adv. Funct. Mater.,* **12**, 323–331.

Yoffe A D (2001), Semiconductor quantum dots and related systems: Electronic, optical, luminescence and related properties of low dimensional systems, *Adv. Phys.,* **50**, 1–208.

Yu G H and Lieber C M (2010), Assembly and integration of semiconductor nanowires for functional nanosystems, *Pure Appl. Chem.,* **82**, 2295–2314.

Zhang J, Lin A, Patil N, Wei H, Wei L, Wong H S P and Mitra S (2012a) Robust digital VLSI using carbon nanotubes, *IEEE Trans. Computer-Aided Design Integr Circuits Sys.,* **31**, 453–471.

Zhang W and Chou S Y (2003), Fabrication of 60-nm transistors on 4-in. wafer using nanoimprint at all lithography levels, *Appl. Phys. Lett.,* **83**, 1632–164.

Zhang Z Y, Wang S and Peng L M (2012b) High-performance doping-free carbon-nanotube-based CMOS devices and integrated circuits, *Chin. Sci. Bull.,* **57**, 135–148.

13

Microfluidics: technologies and applications

L. JIANG and N. S. KORIVI, Tuskegee University, USA

DOI: 10.1533/9780857098757.424

Abstract: This chapter provides an overview of the area of microfluidics in the context of microdevices and systems. The chapter begins with a historical perspective and an introduction to major microfluidic schemes based on fluidic mechanisms. It is followed by a concise review of the present state of technology, including fabrication technologies, interconnect approaches, and mainstream applications such as chemical and environmental, energy, biological and biomedical. Towards the end of the chapter, a short commentary discusses likely trends, challenges, and other aspects of consideration for the future, primarily from engineering and technology perspectives.

Key words: microfluidics, droplet microfluidics, micromolding, PDMS, microfluidic interconnect.

13.1 Introduction

Microfluidics deals with the behavior, control, and manipulation of fluids geometrically restricted to sub-millimeter dimensions, involving the integration of fluids with micro- and nano-structures and devices.[1] It is an interdisciplinary field, overlapping with the areas of microtechnology, engineering, biotechnology, chemistry, physics, and materials science. Consequently, this field comprises a wide variety of applications in different areas. In recent years, the field of microfluidics has grown rapidly with the continuing development of micro-miniaturized analytical devices for biomedical analysis, chemical sensing, genetic analysis, and metabolic monitoring/detection.[2] The advantages of such miniaturized devices include small reagent volumes, fast response times, low cost, disposability, and reduction or elimination of cross-contamination.[3]

Microfluidic devices typically incorporate fluidic channels with at least one dimension in the sub-millimeter scale. These fluidic channels offer high surface-to-volume ratio and consume small volumes of reagents. These characteristics offer advantages for applications such as (bio)chemical analysis. The high surface-to-volume ratio allows a high rate of heat and mass transfer, enabling diverse applications for these devices, such as heat exchange modules for high-power electronics.[4]

After its beginning more than two decades ago, the field of microfluidics found initial application in ink jet printer nozzles and gas chromatography.[5,6] These early devices were developed by traditional silicon micromechanics fabrication methods. Later applications ranged from single component, such as flow sensors and valves, to more complicated integrated microfluidic systems for chemical analysis. Subsequently, microfluidics evolved over the years with new applications facilitated by new fabrication methods. Since its early applications, microfluidics has found many other applications, including some mainstream ones. Some subsequent major applications that have become well-established include microfluidic devices and chips to analyze deoxyribonucleic acid (DNA).

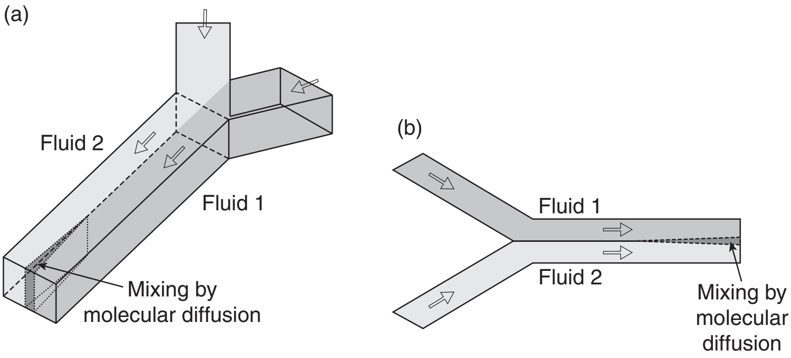
13.2 Current trends in microfluidics

This section will provide a primer to the present major subcategories of microfluidics, based on the fluidic mechanisms employed.

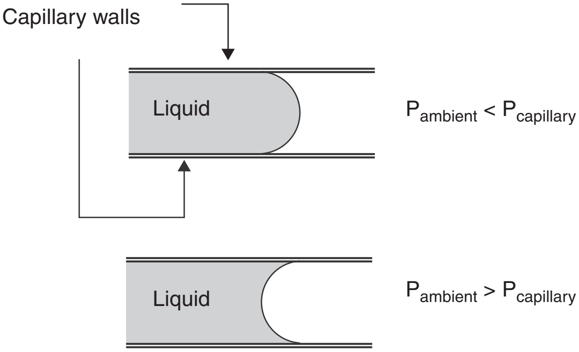
13.2.1 Continuous flow microfluidics

Continuous flow microfluidics, as the name suggests, deals with the manipulation of continuous flow of fluids in the sub-millimeter dimension. Continuous flow devices have existed since the advent of the microfluidics field and have constituted a major part of microfluidics-based applications. An advantage of continuous flow microfluidic devices is the high degree of control of flow characteristics obtained with them. It is also easier to implement these devices for some applications compared to droplet-based devices. Continuous flow microfluidic devices have found a diverse range of applications, including bioanalytical, chemical, energy, and environmental, among others.

A major characteristic of microfluidics is the significance of viscous forces, when compared to fluid flow at the macro-scale. A dimensionless number called the Reynolds number is indicative of the ratio between inertial forces to viscous forces. This Reynolds number in microfluidics is low compared to that in macro-scale fluidics.[7] Consequently, two fluids flowing in parallel streams do not mix easily, with molecules of one fluid permeating into the other fluid by means of diffusion.[8] This phenomenon of microfluidics, known as laminar flow (Fig. 13.1), has defined the way in which several continuous flow microfluidics applications have been designed. In several applications, the laminar flow aspect has necessitated fluidic mixing by means of micro-miniaturized mixer devices. The phenomenon has been employed constructively by continuous flow microfluidic devices and systems for the creation of several novel applications.[9]


(a)
Fluid 2
Fluid 1
Mixing by molecular diffusion
(b)
Fluid 1
Fluid 2
Mixing by molecular diffusion

13.1 (a) Schematic illustration of the laminar flow concept in a microfluidic channel. Two inlet fluidic streams meet and flow parallel to each other as they do not mix easily. (b) A two-dimensional representation of the same schematic. The diagram is not to scale.


Capillary walls
Liquid
$P_{ambient} < P_{capillary}$
Liquid
$P_{ambient} > P_{capillary}$

13.2 Schematic of the capillary effect showing the difference between pressures in a capillary ($P_{capillary}$) and that of the ambient ($P_{ambient}$). For liquid in a small capillary, positive and negative pressure differences arise from a positive or negative curvature of air–liquid interface respectively. The diagram is not to scale.

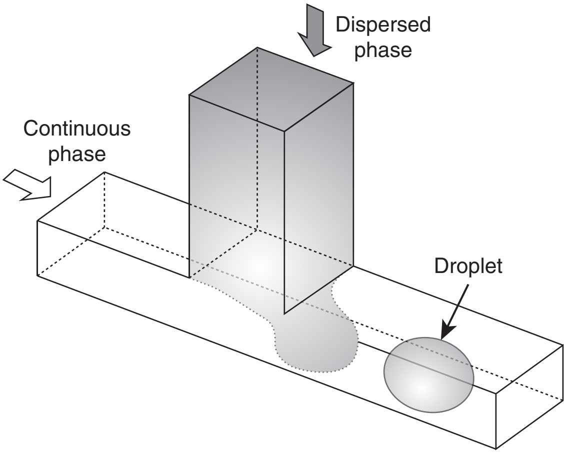
An important aspect of continuous flow microfluidics is the method of liquid propulsion employed. Some of the major propulsion mechanisms are pressure driven, capillary, electrokinetic, centrifugal, and acoustic, among others.[10] In pressure-driven flow, the liquid transport is by pressure differentials in microchannels, often created by external or internal pressure sources. In the capillary mechanism, liquids are driven on a microstructured surface by capillary forces, with liquid movement modulated by the wettability and features of the surface. When liquid is restricted within a small capillary or micro-channel (Fig. 13.2), the liquid front or the air–liquid interface is curved proportionally to its angles of contact with the walls and size of the channel.[11] This curvature causes pressure, which can either prevent liquid from filling the channel (positive pressure, as in the case of a non-wettable channel) or draw the liquid into the micro-channel (negative pressure, as in the case of a wettable channel).

The electrokinetics flow mechanism involves the use of electric fields to control fluidic operations. Subcategories include electrophoresis, dielectrophoresis, and electroosmosis. In the centrifugal mechanism, liquid transport is by centrifugal force, Coriolis force, and capillary force, among others.[12] Liquid operations are typically arranged extending radially outwards from a central position on a rotatable microstructured substrate. The acoustic mechanism involves acoustic shock waves traveling on a hydrophobic solid surface where droplets are located. These waves incite a stream on the solid–liquid interface, causing droplet motion.[13]

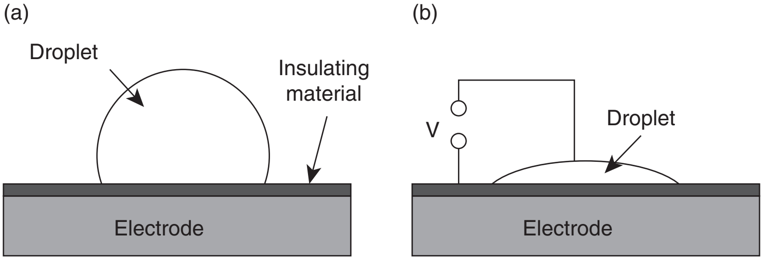
13.2.2 Droplet microfluidics

Droplet microfluidics is a subcategory of microfluidics. Droplet microfluidics, referred to as digital microfluidics, involves the generation and manipulation of discrete volumes of fluids inside microdevices.[14] In contrast to continuous flow devices, droplet microfluidics relies on creating discrete volumes by means of immiscible phases. These discrete volumes or droplets are largely monodisperse with diameters in the nano- and micro-meter range.[15] Since each droplet can be independently controlled, it can function as a microreactor that can be individually controlled, transported, and undergo other fluidic manipulations and analysis.[16] The ability to form and control multiple identical droplets or microreactors allows for a high number of parallel analyses to be performed and large data sets to be collected. This is an important characteristic of droplet microfluidic devices, which makes it easier to scale up these devices in comparison to continuous flow devices. Droplet microfluidics has found several applications in the life sciences and chemistry areas, including DNA and protein analysis.[17]

Droplets can be generated by several different methods, such as those based on microfluidic channel geometry, electrical control among others. Geometry-based methods include microfluidic flow focusing and the use of a T-junction. The flow-focusing method involves the forced flow of a dispersed fluid and a continuous phase fluid through a narrow microfluidic structure, employing symmetric shearing of the dispersed phase by the continuous phase to generate droplets.[18] The T-junction method involves two immiscible fluids being brought together by means of a T-junction configuration of microfluidic channels with the inlet channel containing a dispersed phase fluid orthogonally intersecting a main channel containing a continuous phase fluid (Fig. 13.3).[19] Water is an example of a dispersed phase fluid and of a fluid to be dispersed. Oil is an example of a continuous phase fluid.

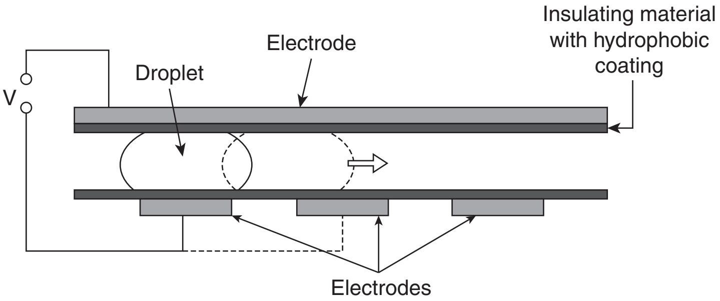


13.3 Schematic of droplet generation by a T-junction consisting of microfluidic rectangular channels. The continuous phase fluid flows in the main channel and the fluid to be dispersed is supplied by the inlet. The diagram is not to scale.



13.4 Schematic representation of the electrowetting principle. (a) A liquid droplet on a hydrophobic surface assumes a natural spherical shape. (b) An applied electric potential or voltage between the droplet and an electrode causes a reduction in the contact angle between the droplet and surface. This causes the droplet to flatten and form a larger footprint on the surface. The diagram is not to scale.

The manipulation of droplets consists of several subcategories, including general movement, mixing, fission, and fusion. One approach to mix, merge, and sort droplets is by geometry-based pressure modulation.[20,21] Another approach is by electrical control techniques, including electrowetting.[22] Electrowetting on dielectric is an effective method to cut, merge, create, and transport liquid droplets.[23] The use of techniques such as those based on electrowetting to move droplets can eliminate the need for microfluidic components such as pumps and valves. Electrowetting is based on controlling the wettability of a surface by means of an applied



13.5 Schematic representation of the electrowetting on dielectric method for droplet actuation or motion. For droplet motion from left to right, an electric potential is applied between a droplet and an electrode at the bottom. Subsequently, this potential is removed and re-applied between an adjacent electrode and the droplet, shown by dotted lines. This causes the droplet to move from the first electrode toward the adjacent electrode. The diagram is not to scale.

voltage (Fig. 13.4). Aqueous droplets attain a natural spherical bead shape on hydrophobic surfaces. When a voltage is applied between a droplet and an insulated electrode, the droplet can spread on the electrode surface.[17] This is due to a decrease in the contact angle between the droplet and the surface.

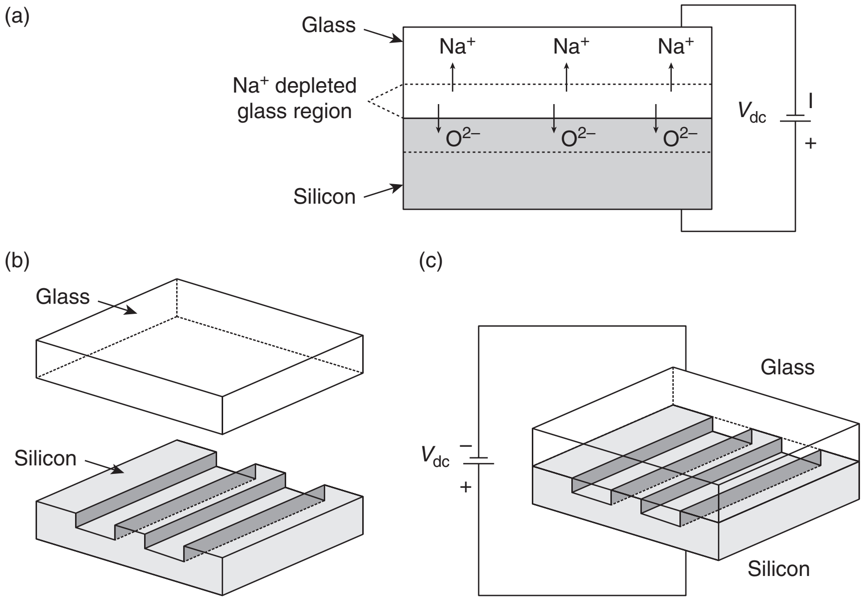
For droplet actuation or manipulation based on the electrowetting on dielectric method, electrical signals to a series of electrodes placed adjacent to each other are employed to define the dimensions and position of droplets. To illustrate, the sequential removal and application of electrical signals to adjoining electrodes can transport droplets between the electrodes (Fig. 13.5).

13.3 Present state of technology

This section reviews the present state of technology in microfluidics, with focus on two major areas: fabrication methods and interconnect technology.

13.3.1 Fabrication methodologies

The fabrication of microfluidic devices is an important factor influencing how these devices and technology are applied. Since the inception of microfluidics, its fabrication technologies have steadily evolved. This section describes some of the fabrication methods used to realize present day microfluidic structures and devices.

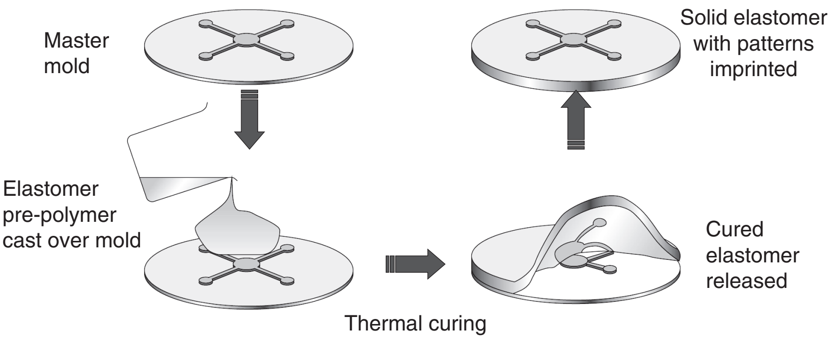


13.6 (a) Schematic representation of anodic bonding of silicon and glass substrates. (b), (c) For microfluidic device fabrication by anodic bonding, the silicon substrate typically consists of micromachined fluidic channels. The diagram is not to scale.

Conventional

Microfluidic devices can be fabricated by several conventional methods using a broad range of materials, such as silicon, glass, plastics, metals, and other materials. These methods include standard integrated circuit (IC) processes and micromachining processes such as bulk and surface micromachining, wafer bonding, LIGA (lithographie galvanoformung abformung or lithography + electroplating + molding), micromolding, and deep silicon reactive ion etching (RIE), among others.

One of the earliest conventional fabrication methods was anodic bonding, which typically bonded silicon to glass surfaces to make microfluidic structures (Fig. 13.6). In anodic bonding, an external voltage is applied to silicon and glass under elevated temperature. A cathode is in contact with the glass and an anode is applied to the semiconductor. At bonding temperature, sodium oxide (NaO_2) in glass dissociates into positively charged sodium ions (Na^+) and negatively charged oxygen ions. The Na^+ ions migrate toward the cathode and a negatively charged depletion layer is formed close to the anode. The large electrostatic force induced by the double charge layer pulls the glass into intimate contact with silicon and a strong bond is formed after a period of time. Anodic bonding is typically employed in conjunction with

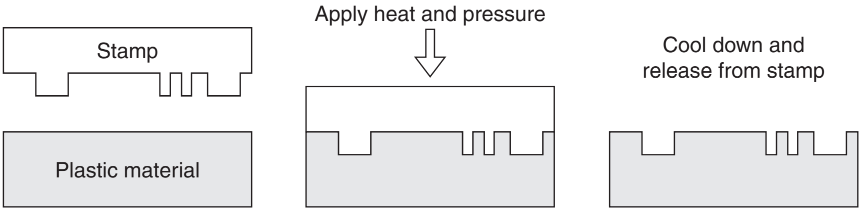


13.7 Schematic diagram of a micromolding process involving elastomers such as polydimethylsiloxane (PDMS). The diagram is not to scale.

surface or bulk micromachining, methods which form microfluidic channels and structures in silicon, glass and other substrates.

Although anodic bonding of silicon and glass was one of the earliest approaches to fabricate microfluidic devices, it is still being used for state-of-the-art submicron or nanofluidic devices due to its merits of hermeticity and strength to withstand high internal pressure.[24] Complete glass chips with micro- and nanochannels are also being developed by standard ultraviolet light photolithography and wet etching technique on glass substrates, sealed to cover glass plates. These fabrication methods are relatively simple and cost-effective and yield glass chips with good visibility.

In the recent years, the use of new materials such as polymers, ceramics, and others has enabled other fabrication techniques including micromolding, soft lithography, micro-stereolithography, and variants of photolithography. In the micromolding method, an elastomeric material in liquid form is poured on a mold stamp with microfabricated features (Fig. 13.7). Subsequent curing of the elastomer over the stamp hardens the elastomer while transferring complementary features from the stamp to the elastomer. The stamp is usually of a hard material such as silicon, metal, or plastic. The features on the stamp are made by micromachining processes. A popular elastomer in microfluidics is polydimethylsiloxane (PDMS), which is a silicone elastomer. When a flexible material such as silicone elastomer is used for the stamp, the method is known as soft lithography.[25] This silicone stamp has features patterned on its surface, and is employed to print complementary structures with feature sizes ranging from nanoscale to a few hundred microns. Soft lithography incorporates a wide variety of techniques organized into three main categories: replica molding, embossing, and printing,[26] all of which include subcategories. The direct printing process is an easy approach to fabricate disposable microfluidic devices.[27]



13.8 Schematic representation of a hot embossing process. The diagram is not to scale.

Hot embossing is another method to make microfluidic devices and chips out of plastics including polycarbonate and poly(methyl methacrylate) (PMMA).[28] This method involves a stamp or master being placed in contact with a plastic surface, followed by the uniform application of pressure while heating the plastic substrate to beyond its glass transition temperature (Fig. 13.8). Subsequently, the plastic substrate is cooled down and the stamp removed. The stamp or master can be made of a hard or soft material, depending on the plastic material to be embossed and the application at hand.

In the technique of micro-stereolithography, prototype three-dimensional models and structures are fabricated by the formation of successive layers.[29] Each layer is formed from liquid photopolymer resin cured by exposure to ultraviolet light. This method is generally appropriate for fabricating arbitrary three-dimensional micro-structures built from carefully-designed two-dimensional slices.

The recent years have witnessed considerable interest in nanofluidics. Several methods have been developed to fabricate structures and devices in the nanoscale dimensions. These include a combination of interferometric lithography and anisotropic etching for fabricating large areas of parallel nanofluidic channels.[30] Focused-ion-beam milling on silicon nitride surfaces has been employed to make precise nanofluidic channel arrays.[31] An electron beam lithography approach for nanofluidic channel fabrication has employed a bilayer resist and thermal reflowing of a top layer of PMMA to seal channels in a bottom layer of photoresist.[32] Among other methods, atomic force microscopy-based lithography has been explored as a nanomachining technique for making long nanochannels on silicon dioxide surfaces.[33]

Non-conventional

Certain applications, including some lab-on-chip schemes, require devices to be disposable and inexpensive. Toward this, new non-conventional methods enabling rapid prototyping and low-cost fabrication are being developed.

These innovations are primarily based on ultra-low-cost substrate materials and rapid-patterning techniques. Paper is one of the substrate materials being explored for bioassays because it is relatively inexpensive, easily available, biodegradable, and easy to functionalize.[34] Paper substrate employed with appropriate photolithography techniques can result in paper-based microfluidic devices. Three-dimensional (3D) microfluidic devices can be fabricated by stacking multiple layers of patterned paper using appropriate adhesive materials.

Interest in inexpensive and disposable devices has resulted in the expansion of boundaries in terms of materials used for fabrication. Numerous non-conventional materials are being investigated for making low-cost and rapidly prototyped microfluidic devices. In the recent years, a diverse range of materials, ranging from polyolefin to pressure sensitive adhesive tape, have been employed to fabricate functional microfluidic structures and components.[35,36]

13.3.2 Interconnect approaches

A major topic of interest while discussing microfluidics is the creation of fluidic interconnections between devices and the external world. These interconnections are often referred to as 'macro-to-micro' interfaces or 'world-to-chip' interfaces. The fluidic interconnection aspect is of significant importance in microfluidics technology, and often defines the reliability of a particular microfluidic approach. The implementation of a successful world-to-chip fluidic interconnect approach is one of the major challenges in the development and adoption of microfluidic devices. One reason is that the need to handle micro-scale fluidic flow is coupled with the diverse nature of microfluidic devices.

A microfluidic device or system typically requires an interface between the micro-scale fluidic structures and the macroscopic world. The dimensions of these structures are in the micro-scale, which can pose challenges when interfacing with the larger dimensions of the external world. Designing a fluidic interconnect approach becomes more complex when a microfluidic device has electrical components. In this case, liquids have to be handled on the same chip as electrical lines. For applications requiring the supply of fluids and reagents to a microfluidic device from the external world, it is a challenge to establish direct interfaces because of the dimensional incompatibilities. There is also the aspect of physical stresses caused in the microfluidic device due to macro-scale fluidic connectors.

An ideal microfluidic interconnect approach is one that is relatively robust and convenient to deploy within a practical set-up time. It should provide high reliability in eliminating fluidic leakages at the operational conditions anticipated. To date, several approaches have been proposed for the packaging and interconnection of microfluidic devices.[37,38] Some of the approaches involve the formation of permanent interconnects that are relatively simple to implement for a low number of interconnections. However, the permanent interconnect approach has drawbacks. To illustrate, some permanent interconnects were implemented by using epoxy to attach fluidic tubing to the device packaging.[39] The drawbacks with this approach are the relatively large dead volumes, permanent tube attachment, and the possibility of channel clogging due to the epoxy. Additionally, fluidic tubing is typically made of plastic that may not adhere well to epoxy, requiring additional time for assembly. Further, the permanent method is not practical in terms of labor when a large number of interconnections are required.

The reversible interconnect approach addresses some drawbacks of the permanent approach, namely in terms of reusability, low dead volumes, and ease of implementation for a larger number of interconnects. Some reversible interconnect methods include those based on machining the tip of fluidic tubing to interlink with an O-ring, or custom-made ring or coupler, to form a sealed connection.[40] Other reversible approaches include press-fit type connections involving placement of a larger element into a smaller recess.[41]

While discussing the general area of micro-to-macro interconnects for microfluidic devices, it is essential to address the aspect of interface standardization. There is a need to develop a standard or generic approach to implement interconnects for microfluidic devices. A standardized interconnect approach can decrease the development time and the associated costs. This aspect is addressed further in Section 13.5.1.

13.4 Applications

The field of microfluidics has found many applications in wide-ranging areas. This section reviews the present uses of microfluidic phenomena, devices and technology, focusing on chemical, environmental, biological, biomedical, energy and other emerging applications.

13.4.1 Chemical and environmental applications

Recent years have witnessed considerable growth of microfluidic applications in monitoring and analysing the environment. This is because of the advantages offered by microfluidic devices, including rapid analysis, portability, and low sample and reagent volumes. Microfluidic devices also allow for the incorporation of multiple processes, such as sample pre-treatment, pre-concentration, separation, and detection on a single platform.[42] This is advantageous in applications such as trace analysis of materials, because the risk of contamination is decreased due to sample transportation processes being precluded. Several microfluidic devices and systems have been developed that employ detection and analysis methods based on electrochemistry, chemiluminescence, laser-induced fluorescence, absorbance, and surface-enhanced Raman spectroscopy, among others. Microfluidic devices have employed enzymatic assays to detect mercury.[43] Microfluidic devices based on methods such as potentiometry and fluorescence have been used to detect multiple ions and heavy metals.[44,45] Recently, microfluidic structures have been integrated with magnetic components to form magnetic bead-based assays to detect bacteria such as methicillin-resistant *Staphylococcus aureus*.[46] Microfluidic devices are being used for chemotaxis of microbes and bacteria.[47,48] More recently, there has been a trend toward cheap and disposable microfluidic devices based on paper for trace detection and analysis of materials.[49,50]

13.4.2 Energy applications

In recent years, microfluidics has found applications in energy, primarily in microfluidic fuel cells. In these energy devices, functions such as fluid delivery, removal, reaction sites, and electrode components are confined to a microfluidic channel.[51] Microfluidic fuel cells can operate in a co-laminar flow mode without a physical barrier to separate the fuel and oxidant species. This is in contrast to conventional fuel cells, which employ a physical barrier such as a proton exchange membrane for the same purpose. There are drawbacks associated with the use of such barriers in fuel cells. This is a major advantage of microfluidic fuel cells. In terms of applications, microfluidic fuel cells can be used to power microsystems, consumer electronics, and for the purpose of generating on-chip power for certain applications. Microfluidic fuel cells have been developed by exploiting fluid flow properties in the micro-scale, including laminar flow and multi-phase flow. There has also been some development in integrating microfluidic structures with enzymes and microbes to generate electricity by bio-electrocatalysis.

In the future, it is possible that microfluidic fuel cells may be used to supply on-chip power for devices and systems in autonomous applications such as remote sensors for environmental security. There may also be some interest in employing these devices as self-sustaining energy sources. The drive toward practical future applications for these fuel cells may require further research and development into passive-feed fuel supply mechanisms, and methods to use fuel from the ambient or nature. This may require the use of various microfluidic mechanisms and other phenomena occurring in nature. Also, this may call for new research in the area of materials and surface modification methodologies that can exploit constructive fluidic phenomena such as osmosis.

13.4.3 Biological and biomedical applications

Since the inception of microfluidic devices and systems, the main driving force for their continued development is their applicability in biological and biomedical applications. The low reagent volumes, and ability to perform multiple processes simultaneously, make these devices suitable for biochemical analysis, high-throughput screening, and high-sensitivity analysis. Microfluidic devices and systems have demonstrated considerable potential in a variety of biological applications, including enzymatic assays, DNA hybridization reactions, and biomolecular separations, among others. In recent years, microfluidics has found increasing usage in the medical area, especially in diagnostic applications. Microfluidic devices allow portability, making them highly viable candidates for point-of-care applications. Microfluidics is now being employed in applications including drug delivery, cancer studies, cell studies, and screening of newborns.

13.4.4 Emerging applications

Microfluidic devices have been employed to synthesize materials and structures in the micro- and nanoscale dimensions. To illustrate, microfluidic devices have been used for fabricating polymeric microspheres by combining phase separation phenomena with fluid flow. Optofluidics is another area where microfluidics is finding new applications. This area is based on the integration of microfluidics with optics. An example of an optofluidic device is one in which microfluidic channels are incorporated with photonic structures. Optofluidics is finding several new applications in environmental monitoring, medical diagnostics, and chemical sensing, among others. More recently, droplet-based microfluidic devices are being explored for thermal management in ICs. The increasing device complexity and processing power in ICs have given rise to problems concerning on-chip cooling.

13.5 Future trends

Any narrative reviewing microfluidics technology should include a discussion on the future of this field. Technological challenges are among the primary challenges facing the future development of microfluidics and application of microfluidic devices. These are closely linked to the commercialization prospects of microfluidic devices. This section discusses these aspects in detail.

13.5.1 Technology challenges

Recent developments point toward an increasing trend and interest in fabrication methodologies that are relatively inexpensive, easy-to-implement, and have shorter turn-around times. This can be expected, because academic research is increasingly being done with an eye toward commercialization. Another factor is the application aspect. We are now seeing the possibility of microfluidic devices being used in field applications in unexpected areas of the world, such as those in the developing countries. In the future, this factor may play an increasing role in influencing the technological and economic aspects of microfluidics. Several challenges need to be addressed in the microfluidics field to realize its inherent benefits and growth potential. This section will discuss these challenges.

Interconnect technology

A major factor impeding the use of microfluidics at the industrial and consumer level is the lack of standardized interconnects. In microfluidics, interconnects allow connection between chips and systems, and also to the outside world. Similar to the role played by standardized interconnects in electronics, microfluidic interconnects will determine the extent to which microfluidics are adopted in general. We see this area as the single biggest challenge to the growth and further commercialization of microfluidics. Numerous interconnect methodologies and approaches have been put forth. However, many of these efforts appear disparate and mostly geared toward a particular application or device platform. Some commercial interconnect approaches exist, but they are not applicable to all situations, primarily because of the restrictions they might place on device material, geometry, and dimensions. In future, efforts must be geared toward finding a universal solution to the interconnect problem. Considering the diverse requirements with regard to material compatibility, flow rates, and pressure limits, it will be challenging to develop a single interconnection approach to address different applications such as biological and chemical analysis. It can be expected that future microfluidic devices may be fabricated from new materials and methods for a diverse range of applications, which can potentially have unique sets of criteria. Consequently, custom design and development of interconnect approaches may continue till better alternatives are found. The interconnect problem may also influence how microfluidic devices will be made and deployed in future. One may see increasing applications based on stand-alone microfluidic devices in disposable format, or otherwise, used in conjunction with standard equipment for fluorescence or Raman microscopy/spectroscopy or any other type of analysis instrumentation.

Standardized components

The field of electronics has standardized components such as transistors, resistors, and capacitors, among others. This is not presently the case in microfluidics. The lack of standardized components in microfluidics is a major challenge. Standardization will enable a relatively easy integration of microfluidic components into existing laboratory environments. It will allow for rapid prototyping, and enable an individual to easily put together a device or system to test an idea. It is expected that the future will see significant effort and development to address this aspect. Already, there has been an increasing effort toward standardization of chip-to-world interfaces and external geometries such as microscope slides or Society for Biomolecular Sciences size micro-well plates.[52]

Materials

The future will see growing interest in alternative materials for microfluidics. An important part of future research activity will focus on searching for new materials to be used either at the device or component level. This search for newer materials will mirror the way microfluidics has evolved in the past, from traditional silicon micromachined structures to other materials such as glass and subsequently to PDMS. The use of PDMS at both component and device level has been widespread due to the various desirable attributes of this material. The interest in this material will continue, with further exploration into extending its applications. This portends studies on tailoring the properties of PDMS, which is inherently not suitable for harsh chemical environments. There is also the aspect of PDMS's permeability to other materials, which can be a limiting factor for certain applications. Already, there is work going on with some research groups successfully imparting glass-like properties to PDMS, while preserving its inherent benefits.

Physics behind phenomena

While several fundamental phenomena in microfluidics are well-understood, there are still others that need further understanding. In this regard, it is pertinent to develop robust tools for modeling microfluidic phenomena. As device and application complexity increases, these modeling tools will allow the researcher to understand how a particular device may function under a set of circumstances. This will allow for convenient design optimization prior to fabrication, and contribute to an increased pace of device development and subsequent deployment.

13.5.2 Commercialization

The market for microfluidic products is expected to grow to US$1.9 billion by the year 2012 with the majority of revenue resulting from diagnostic applications.[53] By all accounts, microfluidics has had a relatively slow pace of commercialization. The lack of standardized components and interconnects is a major reason for this slow pace. Among other reasons, it has been surmised that discussions and publications generate excessive user expectations about device costs and procedural cost saving with respect to economy of scale. These expectations are subsequently not satisfied by device or system manufacturers.[54] This is a plausible factor, since academia typically focuses on early-level research and demonstration of proof-of-concept. Academic researchers may not have sufficient awareness of the modalities of commercialization. There may be an incongruity between commercial reality and what academic researchers consider to be potential cost saving from economies of scale. A factor to consider is that microfluidics involves physical products that may require new manufacturing systems and methodologies, which have to be cost-effective to encourage commercialization. Another factor to consider is that microfluidics is not a product, but rather an enabling technology that gives developers a means to achieve additional features and enhanced performance in their products.[52] One can expect an increasing number of systems in various fields, including clinical and point-of-care diagnostics, chemical and biochemical analysis, and drug discovery and development that will contain important elements of functionality provided by a microfluidic component.[52]

There also appears to be a gap between what is being developed in academia and the needs and actions of industry.[55] This can be addressed by increased collaboration, within academia itself, and between academia and commercial entities. There is an encouraging recent trend of interdisciplinary collaboration in microfluidics research in academia. This has to be sustained and even increased. Such collaboration between people or entities contributing expertise in different areas, including device design, microfabrication, chemistry, biotechnology, and materials science, may result in more robust microfluidic devices and systems being developed at a faster pace. A dynamic pace in research and development can foster commercialization. Another factor influencing the pace of microfluidics development and ultimately commercialization is the lack of modularity. Significant efforts are required to address this aspect. An encouraging harbinger of the future is commercial availability of off-the-shelf microfluidic components that enable fast, inexpensive implementation of development steps, consequently decreasing the entrance barriers for utilization of microfluidic technologies.[52]

13.6 Conclusion

Microfluidics technology has developed significantly in the past several years. Microfluidic devices and systems are finding new applications in areas such as biomedicine, pharmaceutics, environmental science, and energy, among others. The microfluidics field offers many benefits as an enabling technology, and therefore is being considered an important tool in product and technology development. In the future, diagnostic applications will be the driving factor in the growth of microfluidics. We may see a trend toward self-contained integrated microfluidic devices and systems. The aspect of commercialization will increasingly shape the dynamics of research and development of microfluidics in academia. To increase the commercialization of microfluidic devices, major challenges will have to be addressed immediately on interconnects, modularity, and development of standardized components, and the aspect of costs.

The future will see an increasing trend toward cost saving by means of inexpensive devices and fabrication methodologies. The drive toward cheaper devices to be used in field applications may also involve the development of inexpensive and miniaturized power sources, potentially in a thin-film format. From a device material perspective, the area of microfluidics has made a transformation from traditional silicon micromachined devices to the present day polymeric devices. We will see significant research in exploration of alternative and cheaper device materials. However, in the near future, due to the existing industrial base for polymeric devices, we can expect that polymers will continue to be the main materials of choice for device manufacturers, especially in terms of economies of scale. Even as we attempt to move toward cheaper devices, questions still remain on other costs, including that for reagents. Another factor is the financial aspect of regulation, i.e. the cost added on as a particular microfluidic technology satisfies regulation for use in a part of the world. Looking ahead, research questions will have to be addressed on understanding and characterizing the interaction of fluid flow, other aspects including surface phenomena and forces, dynamics of scaling down, and interactions at the molecular level.

13.7 Sources of further information and advice

Microfluidics research groups in academia and national laboratories

This is a list of some academic groups involved in microfluidics research.

Whitesides Research Group, Harvard University
http://gmwgroup.harvard.edu/
Kirby Research Group, Cornell University

http://www.kirbyresearch.com/index.cfm/page/index.htm
The Burns Research Group, University of Michigan – Ann Arbor
http://www.engin.umich.edu/dept/che/research/burns/
Stanford Microfluidics Laboratory
http://microfluidics.stanford.edu/index.html
Sandia National Laboratories
http://www.sandia.gov/microfluidics/

Microfluidics companies

This is a list of a few commercial enterprises in the area of microfluidics. It is beyond the scope of this chapter to list all the companies currently dealing with microfluidic technologies (Table 13.1).

Table 13.1 Microfluidics companies

Company name	Website	Products/services
Advanced liquid logic	http://liquid-logic.com/	Lab-on-a-chip microfluidic devices, etc.
Affymetrix	http://www.affymetrix.com/	Parallel genetic assays (DNA microarrays)
Caliper life sciences	http://caliperls.com/	Microfluidics, liquid handling etc.
Daktari diagnostics	http://www.daktaridx.com/	Point-of-care diagnostics
Micronics		Near-patient, *in-vitro* diagnostics

13.8 References

1. M. Zhang, X. Gong and W. Wen, *Electrophoresis*, Vol. **30**, pp. 3116–3123, (2009).
2. H. Anderson and A. van der Berg, *Lab Chip*, Vol. **4**, pp. 98–103, (2004).
3. E. Igata, M. Arundell, H. Morgan and J.M. Cooper, *Lab Chip*, Vol. **2**, pp. 65–69, (2002).
4. P.Y. Paik, K. Chakrabarty and V.K. Pamula, IEEE Design Test Comput., pp. 372–381, (2008).
5. E. Bassous, H.H. Taub and L. Kuhn, *Appl. Phys. Lett.*, Vol. **31**, pp. 135–137, (1977).
6. S.C. Terry, J.H.Jerman and J.B. Angell, *IEEE Trans. Electron Devices*, ED-**26**, pp. 1880–1886, (1979).
7. N. Rott, *Annu. Rev. Fluid Mech.*, Vol. **22**, pp. 1–11, (1990).
8. P. Tabeling, ‘Introduction to Microfluidics,’ Oxford University Press, New York, USA, (2005).
9. M.U. Larsen and N.C. Shapley, *Anal. Chem.*, Vol. **79**, pp. 1947–1953, (2007).

10. D. Mark, S. Haeberle, G. Roth, F. von Stetten and R. Zengerle, *Chem. Soc. Rev.*, Vol. **39**, pp. 1153–1182, (2010).
11. E. Delamarche, A. Bernard, H. Schmid, A. Bietsch, B. Michel and H. Biebuyck, *J. Am. Chem. Soc.*, Vol. **120**, pp. 500–508, (1998).
12. S. Haeberle, T. Brenner, H.P. Schlosser, R. Zengerle and J. Ducree, *Chem. Eng. Technol.*, Vol. **28**, pp. 613–616, (2005).
13. M.K. Tan, J.R. Friend and L.Y. Yeo, *Lab Chip*, Vol. **7**, pp. 618–625, (2007).
14. D. Belder, *Angew. Chem., Int. Ed.*, Vol. **44**, pp. 3521–3522, (2005).
15. I. Kobayashi, K. Uemura and M. Nakajima, *Colloids Surf., A*, Vol. **296**, pp. 285–289, (2007).
16. R.B. Fair, *Microfluid. Nanofluid.*, Vol. **3**, pp. 245–281, (2007).
17. Advanced Liquid Logic, Inc. USA www.liquid-logic.com (Last accessed on 30 August 2012).
18. L. Yobas, S. Martens, W.L. Ong and N. Ranganathan, *Lab Chip*, Vol. **6**, pp. 1073–1079, (2006).
19. T. Thorsen, R.W. Roberts, F.H. Arnold and S.R. Quake, *Phys. Rev. Lett.*, Vol. **86**, pp. 4163–4166, (2001).
20. K. Liu, H. Ding, Y. Chen and X.-Z. Zhao, *Microfluid. Nanofluid.*, Vol. **3**, pp. 239–243, (2007).
21. Y.-C. Tan, Y.L. Ho and A.P. Lee, *Microfluid. Nanofluid.*, Vol. **4**, pp. 343–348, (2008).
22. P. Paik, V.K. Pamula and R.B. Fair, *Lab Chip*, Vol. **3**, pp. 253–259, (2003).
23. M.G. Pollack, A.D. Shenderov and R.B. Fair, *Lab Chip*, Vol. **2**, pp. 96–101, (2002).
24. C. Song and P. Wang, *IEEE Trans. Nanotechnol.*, Vol. **9**, No. 2, (March 2010).
25. Y. Xia and G.M. Whitesides, *Angew. Chem. Int. Ed.*, **37**, pp. 562, (1998).
26. B.D. Gates, Q. Xu, J.C. Love, D.B. Wolfe and G.M. Whitesides, *Annu. Rev. Mater.*, Vol. **34**, pp. 339–372, (2004).
27. H. Yu, Y. Lu, Y.G. Zhou, F.B. Wang, F.Y. He and X.H. Xia, *Lab Chip*, Vol. **8**, pp. 1496–1501, (2008).
28. G.B. Lee, S.C. Chen, G.R. Huang, W.C. Sung and Y.H. Lin, *Sens. Actuators B*, Vol. **75**, pp. 142–148, (2001).
29. V.K. Varadan, X. Jiang and V.V. Varadan, '*Microstereolithography and other Fabrication Techniques for 3D MEMS*,' John Wiley & Sons Ltd., Chichester, UK, (2001).
30. M.J. O'Brien, P. Bisong, L.K. Ista, E.M. Rabinovich, A.L. Garcia, S.S. Sibbett, G.P. Lopez and S.R.J. Brueck, *J. Vac. Sci. Technol. B.*, Vol. **21**, pp. 2941–2945, (2003).
31. K. Wang, S. Yue, L. Wang, A. Jin, C. Gu, P. Wang, H. Wang, X. Xu, Y. Wang and H. Niu, *IEEE Proc.*, Vol. **153**, No. 1, pp. 11–15, (2006).
32. M. Fouad, M. Yavuz and B. Cui, *J. Vac. Sci. Technol. B.*, Vol. **28**, pp. C6/11–16, (2010).
33. Z.Q. Wang, S. Tung, N.D. Jiao and Z.L. Dong, Proceedings of 5th IEEE International Conference on Nano/Micro Engineered and Molecular Systems (NEMS), Xiamen, China, 20–23 Jan (2010).
34. W.K.T. Coltro, D.P. de Jesus, J.A. F. da Silva, C.L. do Lago and E. Carrilho, *Electrophoresis*, Vol. **31**, pp. 2487–2498, (2010).

35. D. Nguyen, J. McLane, V. Lew, J. Pegan and M. Khine, *Biomicrofluidics,* Vol. **5**, 022209, pp. 1–12, (2011).
36. P. Nath, D. Fung, Y.A. Kunde, A. Zeytun, B. Branch and G. Goddard, *Lab Chip*, Vol. **10**, pp. 2286–2291, (2010).
37. B.L. Gray, D. Jaeggi, N.J. Mourlas, B.P. Van Drieenhuizen, K.R. Williams, N.I. Maluf and G.T.A. Kovacs, *Sens. Actuators A*, Vol. **77**, pp. 57–65, (1999).
38. S. Li and S. Chen, *IEEE Trans. Adv. Packaging*, Vol. **26**, pp. 242–247, (2003).
39. A.V. Pattekar and M.V. Kothare, *J. Micromech. Microeng.*, Vol. **13**, pp. 337–345, (2003).
40. J.H. Tsai and L.W. Lin, *J. Micromech. Microeng.*, Vol. **11**, pp. 577–581, (2001).
41. V. Saarela, S. Franssila, S. Tuomikoski, S. Marttila, P. Ostman, T. Sikanen, T. Kotiaho and R. Kostiainen, *Sens. Actuators B*, Vol. **114**, pp. 552–557, (2006).
42. S. Kondapalli, J.T. Connelly, A.J. Baeumner and B.J. Kirby, *Microfluid. Nanofluid.*, Vol. **11**, pp. 537–544, DOI:10.1007/s10404–011–0819–0, (2011).
43. C.K. Prudente, R.S. Sirios and S. Cote, *Anal. Biochem.*, Vol. **404**, pp. 179–185, (2010).
44. W.-Y. Liao, C.-H. Weng, G.-B. Lin and T.-C. Chou, *Lab Chip*, Vol. **6**, pp. 1362–1368, (2006).
45. S. Kou, S.-W. Nam, W. Shumi, M.H. Lee, S.W. Bae, J. Du, J.S. Kim, J.-I. Hong, X. Peng, J. Yoon and S. Park, *Bull. Korean Chem. Soc.*, Vol. **30**, No. 5, pp. 1173–1176, (2009).
46. C.H. Wang, K.Y. Lien, J.J. Wu and G.B. Lee, *Lab Chip*, Vol. **11**, No. 8, pp. 1521–31, (2011).
47. J.R. Seymour, T. Ahmed, Marcos and R. Stocker, *Limnol. Oceanogr. Methods*, Vol. **6**, pp. 477–488, (2008).
48. T. Ahmed, T.S. Shimizu and R. Stocker, *Integrative Biology*, Vol. **2**, pp. 604–629, (2010).
49. Z. Nie, C.A. Nijhuis, J. Gong, X. Chen A. Kumachev, A.W. Martinez, M. Narovlyansky and G.M. Whitesides, *Lab Chip*, Vol. **10**, pp. 477–483, (2010).
50. J. Ho, M.K. Tan, D.B. Go, L.Y. Yeo, J.R. Friend and H.-C. Chang, *Anal. Chem.*, Vol. **83**, pp. 3260–3266, (2011).
51. E. Kjeang, N. Djilali and D. Sinton, *J. Power Sources*, Vol. **186**, pp. 353–369, (2009).
52. H. Becker, 'Microfluidics: A technology coming of age,' European Medical Device Technology, (1 May 2008).
53. 'Emerging markets for microfluidic applications,' YoleDévelopement, www.yole.fr, June (2007).
54. H. Becker, *Lab Chip*, Vol. **9**, pp. 1659–1660, (2009).
55. N. Blow, *Nat. Methods*, Vol. **6**, pp. 682–685, (2009).

14

Modeling of nanolithography processes

A. ISOYAN and L. S. MELVIN III, Synopsys Inc., USA,

DOI: 10.1533/9780857098757.444

Abstract: Mathematical models of semiconductor processes are heavily used within the microelectronics manufacturing industry. They are used to design and understand molecular-level effects, comprehend and develop local patterning processes, and modify the design pattern to better image the desired pattern on a wafer. To accomplish these tasks, different model types are developed including molecular models, bulk models, and compact models. All models and model applications must resolve the interaction between the competing effects of model computation time and model accuracy. This chapter discusses the basics of mathematical models used for optical, EUV, and e-beam lithography modelings.

Key words: modeling, OPC, lithography, compact model, EUV lithography, E-beam lithography.

14.1 Introduction

Mathematical models of semiconductor processes are heavily used within the microelectronics manufacturing industry.[1] They are used to design and understand molecular-level effects, comprehend and develop local patterning processes, and modify the design pattern to better image the desired pattern on a wafer. To accomplish these tasks, different model types are developed including molecular models, bulk models, and compact models. All models and model applications must resolve the interaction between the competing effects of model computation time and model accuracy.

Molecular models are generally used to work with chemistry, materials, and energy interactions associated with the semiconductor manufacturing process. One example of this type of modeling is a photoresist model. These types of models work at the atomic level. They have been used to design self-assembled polymers and molecular-level photoresists by studying the chemical composition of the photoresist materials and prospective compounds to improve imaging performance.

A second model type is the bulk model. A bulk model is used to explore mechanisms such as line formation in a photoresist. A bulk model usually comprises an optical model interacted with photoresist and other materials as well as substrate topography. Sentarus Lithography by Synopsys, and Prolith by KLA-Tencor, are two commercial packages widely used for bulk modeling in semiconductor modeling.

Compact models are a special type of model that trades model fidelity for extreme runtime. Compact models are used in the optical proximity correction (OPC) process. This allows the model form to be used to model every feature on an entire chip for the purposes of OPC. OPC packages are provided by several companies, including the Synopsys Proteus tool and the Mentor Graphic Calibre tool.

Other models, such as thermal and mechanical stress models, are also used in semiconductor manufacturing. However, their applications are normally found in, but are not limited to, the design of semiconductor manufacturing tools.

14.2 Optical lithography modeling

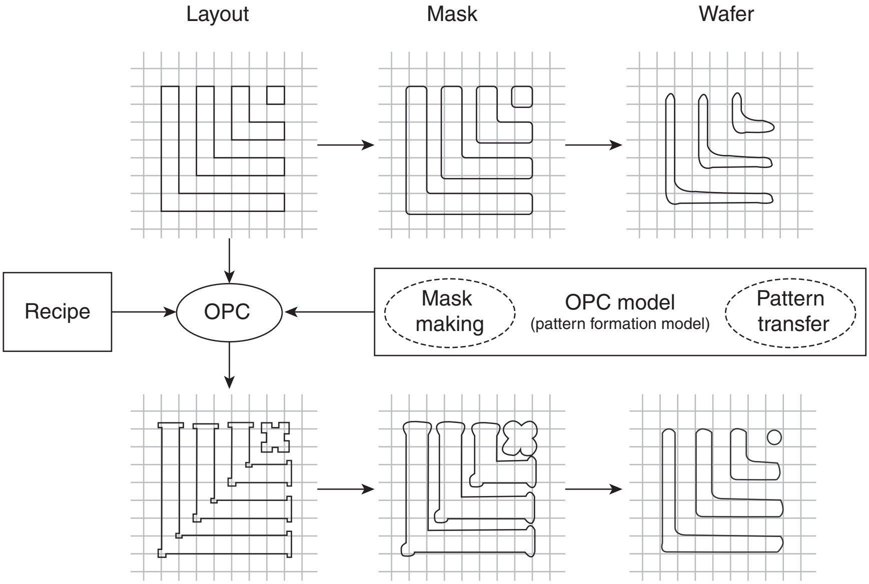
A simple way to improve lithographic resolution is to reduce the wavelength of the imaging system. However, development and manufacturing costs do not always allow wavelength reductions, as evidenced by the use of 193 nm illumination from the 90 nm through the 20 nm half pitch manufacturing node. An alternative strategy for smaller wavelengths and improved illumination is mask correction to compensate for pattern transfer non-idealities at the mask level.[2–4] The approach of making systematic modifications to mask geometries to compensate for pattern transfer non-idealities is typically called OPC.

14.2.1 Optical proximity correction

In general, there are four types of image distortions, namely critical dimension (CD) uniformity, CD linearity, line-end shortening, and corner rounding.[5] These distortions or proximity effects create serious CD control problems in Integrated Circuit (IC) layouts, and hence necessary adjustments to the original mask layout design are needed. The proximity effect origins can be divided into two separate sources: optical, which is caused mainly by the light diffraction,[6] and resist effects.[7] There are also several other sources that introduce imperfections/distortions in the final pattern.[5,8] These additional side effects come mainly from the optical lithography system itself. The optical lithography system, or optical stepper, is a complicated tool and comprises four basic parts: an illumination system, a mask, an exposure system, and a stage that holds the wafers. The illumination system's imperfection, lens quality, and alignment have critical influence on how the image from mask is transferred to the wafer. One should also control the mask quality.[9] Most masks are fabricated using conventional electron beam lithography (EBL), and hence imperfections can be introduced by the fabrication processes.[10] The stage, on other hand, also plays a significant role.[5] Relative vibration between the illumination system and the stage can introduce unwanted image blurring. The stage's vertical location or working distance also should be precisely controlled, since the wafer position should be exactly on the optical system's best possible focal plane, and a small shift out of that plane introduces image quality degradation.[5,8] Thus, overall there are several factors that can potentially have an influence on the quality of the final pattern, and hence corresponding adjustments and corrections are necessary for successful pattern transfer.

The goal of OPC is to compensate for optical, photoresist, and other known effects by making appropriate enhancements to the mask layout. The OPC problem statement can also be formulates as follows: for a given final desired pattern on the wafer, find an input mask design that will generate the desired pattern after all the lithographic steps and photoresist processing. The basic concept of the OPC can be visualized as shown on Fig. 14.1.

Two types of OPC approaches are available for mask modifications: rule-based[5,8,11] and model-based.[8] Rule-based OPC approaches are an extension of the methods used for manual OPC. Through experiment or simulation, specific corrections that should be applied in a given structure are collected. Then, based on known rules and pattern recognition algorithms, appropriate mask modifications are applied whenever that geometrical pattern occurs in the layout. Rule-based OPC approaches are fast by definition, but



14.1 Schematic representation of OPC concept.

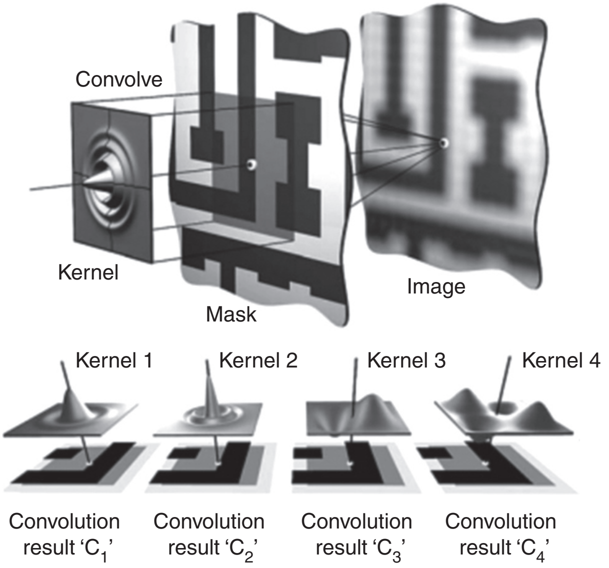
unfortunately they are not as accurate as could be desired. On the other hand, the model-based OPC techniques are completely different from rule-based OPC, in that mathematical models are used to represent the image formation process of the optical lithography system and modify edges on the mask to improve the output pattern fidelity on the wafer. Model-based OPC techniques are more general correction solutions for mask corrections, but can require more OPC time due to intensive calculations and simulations. Depending on how intensive the mathematical models capture the actual lithographic processes, the model-based OPC models can be computationally very expensive, or balanced, or optimized for fast calculation.

14.2.2 Compact models

Compact models are used for OPC problems. In OPC problems, larger areas (1000 of mm^2) must be simulated multiple times to produce a reticle pattern for manufacturing. In a chip with 10^8 transistors, this can be a very significant number of simulations. Each transistor comprises ~1000 polygons and each polygon must be broken up into ~100 pieces; each transistor comprises ~10 layers (one reticle per layer), and each transistor must be simulated ~10 times, for an approximated total number of simulations of 10^{15} to produce a high quality OPC pattern on a reticle set. Assuming each model evaluation can be calculated in a microsecond, the entire reticle set could be simulated in ~10^9 s, or about 32 years.

Thirty-two years is not a very realistic time to wait for an OPC correction; the market window for the device will have been long past. The calculations can be sped up by parallelization, spreading the calculations over 1000 or 10 000 machines, bringing the calculation elapsed time to a figure more on the order of 10 days for 1000 machines or 1 day for 10 000 machines. Of course, this is expensive in terms of machines, but it does bring the run time into a reasonable window. But this also provides for a bare minimum of model usage in an OPC run, which means the full advantage of the model is not realized. In order to use the model both for more problems, and more effectively, assumptions must be made in modeling to reduce the model runtime. This is accomplished by using a compact OPC model.

A compact OPC model is built from two basic inputs: the model form and an empirical data fit. The model form consists of a mathematical description of the photolithographic system such as the optics, photoresist, and mask. Two of the more common optical representations are the Hopkins and Abbe representations,[6] while the most common photoresist model is the Dills model. The model itself comprises two components, intensity at a point, and threshold. With these two items, the shape of a photoresist contour can be found anywhere on a reticle. However, to speed up a compact model,


Convolve
Kernel
Mask
Image
Kernel 1
Kernel 2
Kernel 3
Kernel 4
Convolution result 'C$_1$'
Convolution result 'C$_2$'
Convolution result 'C$_3$'
Convolution result 'C$_4$'

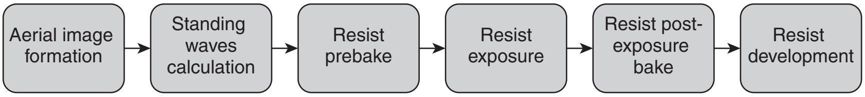
14.2 Convolution kernel-based modeling.

information production is neglected in favor of computational runtime. For example, a compact model is constructed at a specific dose, defocus, and image depth. To vary these parameters, a new model must be constructed, whereas in a full rigorous simulation, these parameters can be varied and the results observed in a short time.

Compact models (Fig. 14.2) comprise a set of kernels that represent the eigenvectors K_i. Each kernel is a table of a set of rapidly calculated rapidly functions that, when interacting with a pattern, can produce the intensity convolution through a series of additions. The additions representation is a method to make the calculation very fast. The output of each kernel interaction with the mask pattern is a coherent system, so the sum of coherent systems is used to achieve the final intensity result:

$$I(x, y) = \sum_{i=0}^{n} c_n \mathrm{Kernel}_n \qquad [14.1]$$

The same speed-up is now achieved with a set of Fourier transforms. In general, the convolution kernel method is faster for sparse patterns and all patterns generally found above the 45 nm half pitch node, while the Fourier transform method is faster for dense patterns. The Fourier transform method and the convolution kernel method produce numerically equivalent results.



14.3 Flow diagram of lithography model.

14.2.3 Image formation modeling

Optical lithography modeling began in the early 1970s to describe the basic steps of the lithography process with mathematical equations.[12–15] These pioneer papers represented the first serious attempt to describe lithography in mathematical form. Lithographic mathematical models simulate the basic steps of all the lithographic processes, starting from aerial image formation, resist exposure, pre- and post-exposure bake (PEB) diffusions, and resist development to obtain the final pattern profile.[8,16] The key needs in such model are to capture most of the problems that can occur during actual real fabrication processes, and to make corresponding corrections to the appropriate fabrication steps. When the systematic component of these errors is characterized, then the OPC system can potentially correct for them and improve the design-to-wafer pattern transfer fidelity. The complexity of OPC models and the number of model components are dramatically increasing for the smaller technology nodes. Figure 14.3 shows a basic schematic of the calculation steps required for lithography modeling.

14.3 The optical system in optical lithography modeling

The optical system of the lithography is one of the most well-studied and well-represented steps, compared to other parts/components of mathematical lithographic models. The extended source method is used to predict the aerial image of a partially coherent diffraction limited or aberrated projection system based on scalar and/or vector diffraction theory.[17] Single wavelength or broadband illumination can be used. The standard image model accounts for the important effect of image defocus through the resist film. The optical system aerial image is modeled using the widely used Hopkins model for general scalar imaging formulation.[18]

The optical system is the first size-limiting step in the entire lithographic process. The main limitation comes from optical diffraction, and the smallest possible minimum feature that can be optically resolved using any optical system is equal to quarter of the illumination wavelength. In general, the optical system's minimum optically resolvable minimum critical dimension (CD_{min}) based on Rayleigh criteria is:

$$\mathrm{CD}_{\min} = k_1 \frac{\lambda}{\mathrm{NA}} \qquad [14.2]$$

where λ is the optical system illumination wavelength, NA is the numerical aperture of the optical system, and k_1 is the optical system's technological representation. Current minimum process dimensions in IC fabrication are in the range of 22–32 nm using 193 nm wavelength immersion optical systems, and k_1 factor somewhere below 0.3.[19] A second important measurement parameter for the optical systems is the depth of focus (DOF).[5] *DOF* is a measure of the adequate DOF where resolvable imaging can still appear, and is defined as follows:[20]

$$\mathrm{DOF} = k_2 \frac{\lambda}{\mathrm{NA}^2} \qquad [14.3]$$

where k_2 is a system's scaling factor. Typically, the k_2 scaling factor for current optical systems is equal to 4. Since the illumination wavelength is more or less fixed, there are only two options left to tweak for improving the minimum resolvable $\mathrm{CD}_{\min}$. The first option is to use optical systems with higher NAs, and the second option is reducing the k_1 factor. However, decreasing the illumination wavelength or increasing NA causes a decrease in the *DOF* and hence it becomes more challenging for successful process development at smaller nodes.

14.3.1 Partially coherent imaging system

In general, widely used optical lithography system illumination sources are partially coherent.[21] In a partially coherent optical lithography system, the light source is located at the focal plane of the condenser lens. Each point of the source emits a coherent, linearly polarized plane wave with spatial frequency defined by the position of the point source relative to the optical axis.[6,21] There are two widely used approaches for the accurate image formation simulation of partially coherent imaging systems, which are Abbe's[22] and Hopkins[18] models. Both approaches are discussed below. In some cases, hybrid Hopkins-Abbe[23] models are also used.

Coherence

Partial coherence is an important and ubiquitous degree of freedom present on most optical lithography imaging systems.[6,24] Partial coherence is defined as the ratio of condenser $\mathrm{NA_C}$ to objective $\mathrm{NA_O}$.

$$\sigma = \frac{\mathrm{NA_C}}{\mathrm{NA_O}} \qquad [14.4]$$

Coherence has a big impact on an image. When $\sigma = 0$ the imaging system is considered coherent, and when $\sigma = 1$ the imaging system is considered incoherent. The k_1 factor can be expressed in terms of the coherence value:

$$k_1 \approx \frac{1}{2(\sigma + 1)} \qquad [14.5]$$

So the $\mathrm{CD_{min}}$ of the imaging system depending on the coherency varies between $0.5\lambda/\mathrm{NA}$ to $0.25\lambda/\mathrm{NA}$. In fully incoherent imaging, the diffraction orders have a spread that causes them to fill the entire pupil. Partial coherence results in an imaging condition where the minimum resolvable pitch is smaller than that possible using coherent imaging; however, modulation of the features decreases linearly to 0. Although the modulation decreases linearly with decreasing feature half pitch, the ultimate resolution of the incoherent system is doubled from that of its coherent counterpart.

14.3.2 Abbe's model

The well-established Abbe formulation is one of today's most common approaches for the accurate image simulation of partial coherent projection systems used in the semiconductor industry. The Abbe's model decomposes the partially coherent imaging system into the superposition of a set of coherent imaging systems.[6,24,22] Each of these coherent imaging systems is based on the contribution of each point within the numerical aperture of the imaging system.[24] The source accepted by the condenser lens is the image of the illumination shape in the lens pupil, or in other words the effective source. Each point of the effective source generates incident plane waves on the object plane. The source is assumed to be incoherent. The mask is located at the object plane. The incident ray emitted from the mask is focused by the imaging lens at the image plane. The wafer is located in the image plane. The intensity of every point in the image plane can be represented as a superposition of all contributions by every effective source point (p_x, p_y) and can be formulated as:[24]

$$I_{\mathrm{image}}(x, y) = \iint_{p_x^2 + p_y^2 \le \mathrm{NA}^2} I_{\mathrm{source}}\left(p_x, p_y\right) E_{\mathrm{image}}\left(x, y; p_x, p_y\right) E^{*}_{image}\left(x, y; p_x, p_y\right) \mathrm{d}p_x, \mathrm{d}p_y \qquad [14.6]$$

where $E_{\text{image}}(x, y;p_x, p_y)$ is the electrical field at image plane produced by an effective source point.

14.3.3 Hopkins diffraction model

Hopkins model is a simplified and approximate version of Abbe's model,[24] where the integration over the source is done before summing up the diffraction angles accepted by the lens.[18] In other words, the Hopkins model is simply a convolution of the source point spread function (PSF) with the mask pattern. This makes the model form extremely compact, hence ideal for OPC usage. To derive the Hopkins diffraction model, $E_{\text{image}}(x, y;p_x, p_y)$ has to be decomposed into two terms. The first term depends only on (p_x, p_y) and represents the effective source. The second term represents the mask, which is located at object plane. After corresponding transformations and derivations, the image intensity distribution delivered at the image plane with a partially coherent illumination system can be expressed as:[18]

$$I_{\text{image}}(x,y)=\int\iiint M(x_1,y_1)T(x-x_1,y-y_1;x-x_2,y-y_2)M*(x_2,y_2)\mathrm{d}x_1\mathrm{d}y_1\mathrm{d}x_2\mathrm{d}y_2 \qquad [14.7]$$

where $M(x, y)$ is the mask transmission function, and $T(x-x_1,y-y_1;x-x_2,y-y_2)$ is the so-called transmission cross-coefficient (TCC), which is given by:

$$T(x-x_1,y-y_1;x-x_2,y-y_2)=J_0(x_1-x_2,y_1-y_2)P(x_1,y_1)P*(x_2,y_2) \qquad [14.8]$$

in which $J(x, y)$ is the mutual intensity function that describes the coherent properties of the illumination system, and $P(x, y)$ the coherent PSF that represents the properties of the projection system. The four dimensional model TCC can be decomposed into a series of its eigenvectors:

$$T(x-x_1,y-y_1;x-x_2,y-y_2)=\sum_{i=1}^{\infty}\lambda_i K_i(x-x_1,y-y_1)K_i^*(x-x_2,y-y_2) \qquad [14.9]$$

Here λ_i is the eigenvalue corresponding to the eigenvector $K(x, y)$. With expansion of a series of eigenvectors and eigenvalues, the image intensity of the Hopkins diffraction model becomes:

$$I_{\text{image}}(x,y)=\sum_{i=1}^{\infty}\lambda_i\left|\iint M(x_1,y_1)K_i(x-x_1,y-y_1)\mathrm{d}x_1\mathrm{d}y_1\right|^2 \qquad [14.10]$$

Each eigenvector $K(x, y)$ can be interpreted as the transfer function of a coherent imaging system. The TCC is independent of the mask pattern. Furthermore, for a given optical system with fixed parameters, the TCC needs to be calculated only once. Thus, the Hopkins diffraction model is more computationally efficient than Abbe's model for a system that varies the input pattern $M(x,y)$.

14.4 Photoresist model

The aerial image intensity distribution calculated by either Abbe's or the Hopkins diffraction models is not enough for a complete representation of the whole lithographic process. Wafers are coated with a light sensitive layer called photoresist.[16] The photoresist reacts to the exposure and its chemical behavior is changed. Exposed photoresists in most cases are developed in the wet etching process. Depending on the photoresist type and developer, during the development process the exposed or unexposed parts of the photoresist are dissolved and cleared out. After the development process, the remaining photoresist forms the printed image on the wafer, which later can be used as an etch mask to transfer that pattern into layers below.

There are positive tone[12] and negative tone[25] photoresists available for commercial use. Positive photoresist is a photosensitive chemical substance that dissolves during the development process in areas where the photoresist has been exposed. On the other hand, negative photoresist dissolves in areas where the photoresist has not been exposed. Depending on the desired pattern, the mask tone, and other further lithographic steps, positive or negative tone photoresist may be used. Photoresists continue to be an active area of research for polymer chemists, as devices of ever smaller size are pursued. Photoresist's main properties are the exposure sensitivity, the resolution, and the line edge roughness.[26] In practice, it is impossible to improve all the three properties at the same time.[27]

The most common photoresists used in industry are chemically amplified photoresists (CAR).[28] CAR photoresists are very sensitive to exposure, which reduces the amount of exposure energy required to produce the chemical change necessary for the development process. In CAR photoresist, exposure light frees a proton (H+) as a catalyst for a deprotection chemical reaction. All positive tone CARs are made up of photoacid generator (PAG) molecules and resist polymers, to which protection groups P are attached.[29] The protection groups strongly influence the solubility of the resist polymers in aqueous solution. The notion of PAG stems from the fact that when a PAG molecule is exposed with light it produces acid (proton) A. The key role of the acid during PEB of a positive tone CAR is to catalyze the removal of hydrophobic protection groups from the polymer so the polymer may be developed away.[30] By diffusing through the resist matrix a small amount of initially generated acid molecules can decompose many protection groups within the resist, effectively amplifying the number of photons used during exposure. Further additives such as base molecules Q (quenchers) may also be present. Quenchers can partially neutralize the acid A and they can also diffuse through the resist. The most significant change after PEB of a CAR is the local change in the protection group concentration, because the protection group concentration strongly affects the resist's solubility during the subsequent development step. However, since the de-protection reaction is catalyzed by acid molecules, and since the local acid concentration can be strongly impacted by the local quencher content, both, the interactions (reaction kinetics) between the concentrations for acid, quencher and protection groups, as well as the diffusion of acid and quenchers, must be treated for modeling the protection group concentration after PEB.

Any simulation approach for photolithography requires the pattern formation process to be separated into at least two main steps: exposure and PEB. During the exposure step, a photosensitive coating (resist) on the wafer is irradiated with a mask specific intensity distribution that causes the solubility of the resist coating to be modified. The chemical modification describing the switch from an insoluble to a soluble state takes places immediately after the exposure and during the PEB.

A serious complication for resist modeling is that accurate physical modeling is computationally expensive.[7] While computational speed is already a tough requirement for rigorous lithography simulators typically dealing with small chip areas of a few square microns, the speed requirements are even more stringent for OPC engines, where the resist model is to be applied for the full chip. In recent years, several approaches for physical resist modeling have been developed to approximate the photoresist development;[8,12,31] however, it is generally difficult to achieve both sufficient levels of accuracy and practical speeds of calculation. Hence it is necessary to model process simulation using specially developed approximation algorithms for low CPU time consumption, calibration accuracy (e.g. measured by the goodness of fit), and the model's predictive power.

14.5 Model critical dimension (CD) extraction

Combined aerial image and resist modeling form a complete model that can be used to simulate the image formation process on the wafer. The next step is to apply a methodology to extract/predict the simulated CDs from the model. This step is necessary for model quality verification, assessment, and predication analysis. Modeled CDs are also used to construct cost functions to drive the model optimization process. Two widely used approaches used for model CD extraction are:[32] constant threshold[31,32] and variable threshold.[33] In the first approximation, threshold models can be associated with the resist development process.

14.5.1 Constant threshold model

The constant threshold model is the simplest approach to extracting the model's simulated CD at specific locations of the formed pattern. The threshold level can be associated with the minimum required exposure energy that clears out the photoresist during the development process. Thus the areas where the exposed energy distribution is greater than the threshold are dissolved, whereas the areas below the threshold level remain. Despite the fact that the constant threshold model has more physical meaning, it is too simplified to be able to describe the entire resist development processes accurately.

14.5.2 Variable threshold model

Variable threshold resist models use a function to determine the threshold to be applied to a specific aerial image to predict CDs. The aerial image is calculated for a small area of a circuit around a feature of interest, and the threshold is applied only locally. The function typically uses features extracted from the aerial image, such as the local image intensity maximum or slope. It is also possible to use features extracted from the pattern being simulated, such as the reticle CD. The function is an algebraic expression of the input parameters with the coefficients of the terms to be determined.[33]

14.6 Difficulties in modeling

14.6.1 Model calibration

One of the problems in OPC model calibration is the determination of the optimized values of all the model parameters involved in the OPC model. OPC models themselves are complex mathematical models, and can contain hundreds of parameters. A straight forward and simple approach for determination of the OPC model parameters is the global search[34] over the entire OPC model to find the model global minimum. The method is well described, and has been used in many applications over the past decades. However, OPC model components, both physical and statistical, can have a certain level of similarity, and in some cases the global optimization method can lead to failed model optimization by degradation of model physicality. Other optimization methodologies are available for model calibration.[35] Regardless of what approach is used for model calibration, experimentally measured data are required. The experimental data are obtained from the actual wafers, later used to train the OPC model. The data from the wafers are collected using scanning electron microscopes (SEM), followed by special algorithms to extract the measured CDs at specific locations on the pattern layout. The quality of the measured data is critical for successful model calibration. Two different types of data points are usually obtained: one-dimensional (1D) and two-dimensional (2D). Typical 1D structures are lines and spaces, and 2Ds are line ends. The collected data, or so-called metrology, along with the coordinates on the layout, are passed to the OPC model optimizer. The optimizer is an engine that, for given possible combinations of certain model parameters, calculates the simulated CDs at specified locations on the layout, compares the modeled and measured CDs, and drives the model calibration in the direction of best possible match with the metrology. The cost function is used to quantify the model's match with the metrology. Various cost functions can be used during model calibration. Typical cost functions are constructed from the root mean square of the fitting errors ($\Delta CD = CD^{mod} - CD^{meas}$), and can be expressed as: $\text{Cost} = \text{RMS}_{\text{All}}\left(\text{RMS}_{1D} + \text{RMS}_{2D}\right)$.

14.6.2 Model runtime and model accuracy

Runtime and accuracy are competing factors in any modeling problem. Model executions that take days or weeks to perform may be too late for the purpose of their execution, while a model that takes minutes or hours to execute may not provide information to an accuracy level that is better than noise.

Some of the larger contributions to runtime:

- Sampling rates – the sampling rate determines the data density in a model. The smaller the sampling rate, the more data. However, the smaller the sampling rate, the more calculations that are needed for the model.
- Asymmetric systems – A symmetric system can be used to reduce the calculation immensely. In a quasar system, the model calculation time can be reduced by almost 4×, because the four sector calculations can be done as one wedge and duplicated. However, the asymmetric system must be calculated for each sector. Approximating patterns to gain symmetry is a technique that can be employed to reduce runtime, but it does result in some accuracy degradation.
- Convolutions – convolutions of optical system kernels with layout are the most expensive computational mathematic functions. They generally cost 10^3–10^5 times more than the least expensive computational function, addition. Reducing and optimizing the number of convolutions in a model can dramatically speed up the model calculation time.
- Pattern area – lithography effects are interference effects, so enough pattern must be modeled to generate the correct diffraction interactions. However, the modeling time is proportional to area, so model time rises

as a square as more patterns are added to the modeled area. This makes the pattern area a significant contributor to model runtime.

Some of the larger contributions to accuracy are:

- Pattern sampling rates – the denser the sampling of input data, the more accurate a model becomes. This is due to less interpolation between model points, as well as better data.
- Model form – the model form may have a significant impact on model accuracy. However, the more complex the model form used, the more convolutions are necessary, leading to increased runtime.
- Pattern area – lithography effects are interference effects, so enough pattern must be modeled to generate the correct diffraction interactions.

To achieve a computable model, these contributions, as well as many others, will have to be balanced to achieve acceptable quality simulation results in an acceptable runtime.

14.7 Extreme ultraviolet (EUV)/electron beam lithography modeling

As the microelectronics industry continues to shrink integrated circuit device sizing, traditional optical lithography is reaching its own resolution limit.[36] Consequently, next-generation lithography candidate technologies are becoming increasingly significant to future process nodes. There are several different candidates for next-generation lithography, such as the extreme ultraviolet (EUV) lithography described in Chapter 2, EBL in Chapter 3, focused ion beam lithography in Chapter 4, and nanoimprint lithography in Chapter 9. Each of these candidates has its own advantages and disadvantages, and there is still uncertainty as to which one will carry the leadership of the next-generation lithography.[37,38] Focused ion beam lithography and nanoimprint lithography are not in the main stream development paths for integrated circuits at the 20 nm half pitch node and below at the time of publication (2012). As such, very little effort has been put into production modeling for these techniques at this time.

14.7.1 E-beam lithography modeling

The proximity effect of electron scattering in resist and substrate causes pattern distortions and printability issues. Applications of EBL in the microelectronics industry require proximity effect correction methods that are both accurate and computationally efficient, in order to maintain CD control and to handle large volumes of data for full chip layouts, especially for future process nodes. The most critical issue in the e-beam simulation work is to understand the interaction between the electrons and resist. Several different models to describe this interaction have been developed previously, based on the PSF convolution using PSF fit models such as multiple Gaussians,[39] analytically derived PSF from transport theory, and the Monte Carlo model. The major disadvantage of the Monte Carlo approach is that it requires significant calculation time in order to obtain a result with sufficiently small statistical fluctuations, especially in the case of a multilayer substrate where a lot of boundary conditions are involved. An alternative to the Monte Carlo method is an analytical representation of the PSF using a Gaussian fit model of electron scattering in resist.[40]

Quick and correct determination of the PSF parameters, such as the forward scattered electrons exposure distribution and the backscattered exposure distribution, has significant importance for aerial image formation, allowing the performing of EBPC effects corrections. Usually for parameter determination, special patterns are exposed, for which the proximity function can be solved analytically, and/or Monte Carlo simulations are used. Both methods require a significant amount of time for experiments, metrology, and computation. Another way to determine the proximity function is to measure it directly, using the point-exposure method of Rishton and Kern.[41] Also, experimental measurement of PSF parameters is influenced by resist development processes and, as a result, this could be one of the reasons for the widespread measured parameters for identical conditions. In general, the forward scattered electrons exposure distribution, the backscattered exposure distribution, and their ratio, are determined by a cross-section of elastic and inelastic scattering of fast electrons in a resist and a substrate, and thus they depend on electron energy E, atomic number, atomic weight, density, resist film thickness, multilayer substrate boundary conditions, etc. Exact values of the parameters are needed for EBL PSF analytical form, and for successful correction of the proximity effect in EBL.

Another fast, and more predictable, method for prediction of EBL PSF optimized parameters for proximity effect corrections is model-based analyses and interpretation of non-corrected representative pattern generic pattern distortions in order to achieve the best possible matching of EBL proximity effects with extracted empirical data. In this case, there is no need for special calculations and simulations, since the EBL model will adjust itself accordingly so that the PSF parameters match the empirical data from the test wafer or actual chip. Compared to other methods, this approach is fast, effective, does not require any additional technology steps and uses only standard measurement techniques for extracting correct measured linewidths values.

E-Beam point spread function

The e-beam lithography system PSF can be approximated in the form of the sum of Gaussian functions, of which the most popular and well-known form is the so-called double Gaussian. Each of the Gaussian distributions describes an effect observed in the EBL, such as the short range (forward scattered electrons) and the long range (backward scattered electrons) energy deposition in the resist. The simplest two-Gaussian PSF can be written as:

$$\mathrm{PSF}_{\mathrm{EBL}}(r) = \frac{1}{\pi(1+\eta)}\left(\frac{1}{\alpha^2}\exp\left(-\frac{r^2}{\alpha^2}\right) + \frac{\eta}{\beta^2}\exp\left(-\frac{r^2}{\beta^2}\right)\right) \qquad [14.11]$$

The first Gaussian distribution describes the deposited energy distribution of the forward scattered electrons in the resist with α standard deviation, and the second one, which is characterized by the parameter β, interprets the backscattered electrons deposited energy in the resist. The backscattered electrons are the primary electrons scattered back from the substrate to the resist. The η characterizes the ratio of the backscattered electrons contribution vs the forward scattered. The PSF*(r)* is the normalized distribution of the deposited energy in the resist by electrons, thus:

$$\int_0^\infty \mathrm{PSF}_{\mathrm{EBL}}(r)\mathrm{d}r = 1 \qquad [14.12]$$

The two-Gaussian approximation model has demonstrated success in predicting the proximity effects in various EBL applications. However, in some conditions the two-Gaussian approximation is not enough to fully describe the EBL model. Adding extra Gaussian terms in the expression for the deposited energy distribution in the resist yields the flexibility to fit the energy distributions, providing more optimal results.[39] It is essential that adding even more Gaussian terms in the PSF will help to adjust more finely the PSF form to meet the experimental data and hence will provide better flexibility for EBL proximity correction.

Electron beam lithography (EBL) compact model

Generally, EBL is a complicated hardware system, and can introduce additional distortions such as aberrations, exposure field distortions, and stage vibrations. These effects can be represented by a single or several kernels, which are applied to the PSF of the ideal EBL.

An aerial image formation can be done by convoluting the EBL PSF with the representative test pattern:

$$I(r) = \iint \mathrm{PSF}_{\mathrm{EBL}}(r) \otimes D(r)\mathrm{d}r$$
$$D(r) = \begin{cases} > 0 & if\,(r) \in P \\ 0 & if\,(r) \notin P \end{cases} \qquad [14.13]$$

where *I(r)* is the deposited electrons energy intensity in the resist at certain resist depth, *P* is the pattern structure such as polygon or line, and *D(r)* depends on the exposure dose assigned to that particular structure/pixel. Unlike optical lithography simulations, where the mask object used in the convolution has constant transmission and phase shift for a given layer, the exposure system employed in EBL optimization could have continuous or discrete transmission values, even for the same layer. EBL is maskless lithography and the pattern generation system is used to create the pattern. Each pixel of the structure can be exposed at a different exposure energy. The EBL compact model allows running fast and accurate full chip process models for EBL.

Similar to optical lithography, deposited electrons energy distribution can be incorporated with resist modeling which allows simulation of the e-beam lithography. The EBL models calibrated using optimization methodologies similar to optical lithographic model calibration.

14.7.2 EUV lithography modeling

Extreme ultraviolet lithography (EUVL) is a leading candidate for the 22 nm node lithography and beyond.[37] EUVL aerial image formation through modeling is necessary for modeling of optical transfer function to assimilate optical diffraction, long range layout dependent flare effects, and shadowing effects due to non-telecentric imaging optics in the EUV case.

Flare in the EUVL optical system is unwanted scattered light that appears beneath dark features, degrading the aerial image contrast, which leads to a reduced size of the process window.[42] The total integrated scatter (TIS) is given by equation:[43]

$$\mathrm{TIS} = \left(\frac{4\pi\sigma_S \cos\theta_i}{\lambda} \right)^2 \qquad [14.14]$$

where σ_s is the standard deviation of the surface roughness, θ_i is the angle of incidence and λ is the wavelength of light illuminating the surface. This indicates that the total flare is inversely proportional to the wavelength squared and proportional to the optical surface roughness squared. The amount of flare is non-uniform across the field of view and is dependent on the density map of the pattern.[44] The amount of flare present can be quantified and characterized by the PSF $\mathrm{PSF}_{\mathrm{SC}}(r)$ due to scattering[45]and the mask reflection factor $T(r)$, and can be expressed as:

$$\mathrm{Flare} = \iint \mathrm{PSF}_{\mathrm{SC}}(r) T(r) \mathrm{d}r \qquad [14.15]$$

This means that by convoluting the density map of the mask design with the $\mathrm{PSF}_{\mathrm{SC}}(r)$, it is possible to calculate a flare map of the design.[44] Non-telecentricity of the reticle (at the image plane) in the EUVL projection optics system causes an asymmetry of the diffracted light coming from the illuminated mask with a three-dimensional topography. This leads to the shadowing of an illuminating beam by an edge of the absorber. This result is a bias of the printed features that generates an overlay error.[45] The bias associated with printed CDs is dependent on both the azimuthal angle (h) changing through the scanner slit and an incident angle (U) accounting for the non-normal illuminating light. Thus, shadowing is dependent on the feature orientation relative to the azimuthal angle (h), where the change in the angle will cause different biases and pattern shifts at different slit positions. In most cases, $\mathrm{PSF}_{SC}(r)$ can be represented as a sum of two Gaussians.

A typical EUV projection setup, an off-axis EUV illumination ray is projected over the mask, and with the aid of reflective optics an image is formed at the image plane.[46] In EUV, the aerial image of a mask, or the light intensity in the image plane of an optical projection system, is modeled using scalar diffraction theory without imposing paraxial approximations. Scalar diffraction theory produces results of adequate accuracy for numerical apertures (NA) ≤ 0.6, and therefore is applied since most of the EUV systems have an NA less than 0.3. EUV aerial image intensity distribution can be defined in the same way as optical lithography using Hopkins equation and the TCC matrix.

Shadowing of an illuminated beam by an edge of the absorber results in a bias of the printed features and generates an overlay error. This is modeled by simulating bias through an offset layer formed with Boolean operations on the original mask layout, then simulating the width of this layer and its transmission coefficient.

Similar to optical lithography, EUV aerial image distribution, along with flare and shadowing compensations, is incorporated with resist modeling. The EUV models are calibrated using optimization methodologies similar to optical lithographic model calibration.

Table 14.1 Modeling technique properties

Model type	Molecular level	Bulk level	Electron path	Area	Optics	Z-Direction stack	Electron exposure
Monte Carlo	Yes	No	Yes	1 μm^3	No	No	No
Rigorous simulation	No	Yes	No	100 μm^2	Yes	Yes	Yes
Compact modeling	No	Yes	No	1000 + mm^2	Yes	No	Yes

14.8 Conclusion

This description discusses the basics of mathematical models used for optical and patterning components of semiconductor manufacturing processes. The rigorous and compact model forms discussed are the basis of process development and process correction for current and future half pitch nodes. Rigorous models are one of the components of process look ahead to reduce the cost of moving to more aggressive process nodes. Compact models are the enabling technology for OPC, the technology that is mandatory for preparing a lithographic reticle for manufacturing. The basic study presented here provides a basis for understanding how to use these modeling technologies. Table 14.1 helps to distinguish features of different modeling techniques presented in this chapter.

The techniques presented here are the most common modeling techniques used in semiconductor manufacturing at the time of printing. Other methods, such as ray tracing or brute force molecular modeling, are also available, but are not as popular in manufacturing.

14.9 References

1. A. R. Neureuther, 'Understanding lithography technology', Electronics Research Laboratory, University of California at Berkeley Memorandum UCB/ERL, p. 93–40, (1993).
2. L. W. Liebmann, B. J. Grenon, S. Schomody, M. A. Lavin and T. Zell, 'Optical proximity correction: a first look at manufacturability', *Proc. SPIE,* **2322**, p. 229, (1994).
3. M. D. Levenson, N. S. Viswanathan and R. A. Simpson, 'Improving resolution in photolithography with a phase-shifting mask', *IEEE Trans. Electron Devices*, vol. **29**, no. 12, pp. 1828–1836, (1982).
4. A. Neureuther, 'If it moves, simulate it!', *Proc. SPIE,* **6924**, p. 692402, (2008).
5. A. K. -K. Wong, *Resolution Enhancement Techniques in Optical Lithography*. Bellingham, Washington: SPIE, (2001).
6. M. Born and E. Wolf, *Principles of Optics*, 7th ed. Cambridge, United Kingdom: University Press, (2006).

7. F. H. Dill, 'The basis for lithographic modeling', *Proc. SPIE,* vol. **5754**, p. 377–382, (2005).
8. C. A. Mack, *Fundamental Principles of Optical Lithography: The Science of Microfabrication*. Chichester, England: John Wiley & Sons Ltd, (2007).
9. M. L. Rieger and J. P. Stirniman, 'Mask fabrication rules for proximity-corrected patterns', *Proc. SPIE,* vol. **2884**, p. 323, (1996).
10. C. Pierrat, J. G. Garofalo, J. DeMarco, S. Vaidya and O. W. Otto, 'Rule-based approach to e-beam and process-induced proximity effect correction for phase-shifting mask fabrication', *Proc. SPIE,* vol. **2194**, p. 298, (1994).
11. R. Shi, Y. Cai, X. Hong, W. Wu and C. Yang, 'Important works about rules in rules-based optical proximity correction', *Chinese J. Semiconductors*, vol. **23**, no. 7, p. 701–706, (2002).
12. F. H. Dill, W. P. Hornberger, P. S. Hauge and J. M. Shaw, 'Characterization of positive photoresist', *IEEE Trans. Electron Devices*, vol. **22**, no. 7, p. 445–452, (1975).
13. F. H. Dill, 'Optical lithography', *IEEE Trans. Electron Devices*, vol. **22**, no. 7, p. 440–444, (1975).
14. F. H. Dill, A. R. Neureuther, J. A. Tuttle and E. J. Walker, 'Modeling projection printing of positive photoresists', *IEEE Trans. Electron Devices*, vol. **22**, no. 7, p. 456–464, (1975).
15. K. L. Konnerth and F. H. Dill, 'In situ measurement of dielectric thickness during etching or developing processes', *IEEE Trans. Electron Devices*, vol. **22**, no. 7, p. 452–456, (1975).
16. K. Suzuki and B. W. Smith, *Microlithography: Science and Technology*. Boca Raton, FL, USA: CRC Press, (2007).
17. C. Mack, 'Comparison of scalar and vector modeling of image formation in photoresist', *Proc. SPIE,* vol. **2440**, p. 381, (1995).
18. H. H. Hopkins, 'On the diffraction theory of optical images', Proceedings of the Royal Society of London. Series A, Mathematical and Physical, vol. **217**, no. 1130, p. 408–432, (1953).
19. B. Fay, 'Advanced optical lithography development, from UV to EUV', *Microelectron. Eng.*, vol. **61–62**, p. 11–24, (2002).
20. M. Rothschild *et al*., 'Photolithography at 193 nm', *The Lincoln Laboratory J.*, vol. **10**, no. 1, p. 19–34, (1997).
21. B. Salik, J. Rosen and A. Yariv, 'Average coherent approximation for partially coherent optical systems,' *J. Opt. Soc. Am. A*, vol. **13**, p. 2086–2090, (2001).
22. P. Evanschitzky, A. Erdmann and T. Fuehner, 'Extended Abbe approach for fast and accurate lithography imaging simulations,' in 25th European Mask and Lithography Conference (EMLC), Dresden, Germany, p. 1–11, (2009).
23. K. Adam and M. C. Lam, 'Hybrid Hopkins-Abbe method for modeling oblique angle mask effects in OPC,' *Proc. SPIE,* vol. **6924**, San Jose, CA, USA, p. 69241E, (2008).
24. X. Ma and G. R. Arce, *Computational Lithography*, 1st ed., G. Boreman, Ed. New York, USA: Wiley Series in Pure and Applied Optics, (2010).
25. J. M. Shaw, J. D. Gelorme, N. C. LaBianca, W. E. Conley and S. J. Holmes, 'Negative photoresists for optical lithography,' *IBM J. Res. Dev.*, vol. **41**, no.1.2, p. 81–94, (1997).

26. J. Shin, G. Han, Y. Ma, K. Moloni and F. Cerrina, 'Resist line edge roughness and aerial image contrast,' *J. Vac. Sci. Technol. B*, vol. **19**, no. 6, p. 2890, (2001).
27. G. M. Schmid *et al.*, 'Resolution limitations in chemically amplified photoresist systems,' *Proc. SPIE,* vol. **5376**, Santa Clara, CA, USA, p. 333, (2004).
28. S. J. Bukofsky *et al.*, 'Imaging of photogenerated acid in a chemically amplified photoresist,' *Appl. Phy. Lett.*, vol. **73**, no. 3, p. 408–410, (1998).
29. G. M. Wallraff *et al.*, 'Sub-50 nm half-pitch imaging with a low activation energy chemically amplified photoresist,' *J. Vac. Sci. Technol. B*, vol. **22**, p. 3479, (2004).
30. G. M. Schmid, M. D. Stewart, V. K. Singh and C. G. Willson, 'Spatial distribution of reaction products in positive tone chemically amplified resists,' *J. Vac. Sci. Technol. B*, vol. **20**, p. 185, (2002).
31. T. A. Brunner and R. A. Ferguson, 'Approximate models for resist processing effects,' *Proc. SPIE,* vol. **2726**, p. 198, (1996).
32. W.-C. Huang *et al.*, 'Two threshold resist models for optical proximity correction,' *Proc. SPIE,* vol. **5377**, Santa Clara, CA, USA, p. 1536, (2004).
33. J. Randall, K. Ronse, T. Marschner, M. Goethals and M. Ercken, 'Variable-threshold resist models for lithography simulation,' *Proc. SPIE,* vol. **3679**, Santa Clara, CA, USA, p. 176, (1999).
34. R. Horst, P. M. Pardalos and N. V. Thoai, *Introduction to Global Optimization*, 2nd ed. Dordrecht, The Netherlands: Kluwer Academic Publishers, (2000).
35. A. Isoyan, J. Li and L. Melvin, 'Stepwise fitting methodology for optical proximity correction modeling,' *Proc. SPIE,* vol. **7640**, p. 764030–764031, (2010).
36. B. J. Lin, 'Optical lithography – present and future challenges,' C. R. *Physique*, vol. **7**, p. 858–874, (2006).
37. B. J. Lin, 'Marching of the microlithography horses: electron, ion, and photon: past, present, future,' *Proc. SPIE,* vol. **6520**, p. 652002, (2007).
38. D. Bratton, D. Yang, J. Dai and C. K. Ober, 'Recent progress in high resolution lithography,' *Polymers for Advanced Technologies*, Special issue on lithography, vol. **17**, no. 2, p. 94, (2006).
39. S. J. Wind, M. G. Roseneld, G. Pepper, W. W. Molzen and P.D.Gerber, 'Proximity correction for electron beam lithography using a three-Gaussian model of the electron energy distribution,' *J. Vac. Sci. Technol. B*, vol. **7**, no. 6, p. 1507–1512, (1989).
40. T. H. P. Chang, 'Proximity effect in electron beam lithography,' *J. Vac. Sci. Technol. B*, vol. **12**, p. 1271, (1975).
41. S. A. Rishton and D. P. Kern, 'Point exposure distribution measurements for proximity correction in electron beam lithography on a sub-100 nm scale,' *J. Vac. Sci. Technol. B*, vol. **5**, no. 6, p. 135, (1987).
42. M. Chandhok, S. H. Lee and T. Bacuita, 'Effects of flare in extreme ultraviolet lithography: learning from the engineering test stand,' *J. Vac. Sci. Technol. B*, vol. **22**, no. 6, p. 2966, (2004).
43. J. C. Stover, *Optical Scattering; Measurement and Analysis*, 2nd ed. Bellingham, WA: SPIE Optical Engineering Press, (1995).
44. G. F. Lorusso, A. M. Goethals, R. Jonckheere, J. Hermans and K. Ronse, 'Extreme ultraviolet lithography at IMEC: shadowing compensation and flare mitigation strategy,' *J. Vac. Sci. Technol. B*, vol. **25**, no. 6, p. 2127, (2007).

45. C. Krautschik, M. Ito, I. Nishiyama and S. Okazaki, 'Impact of EUV light scatter on CD control as a result of mask density changes,' *Proc. SPIE,* **4688**, p. 289, (2002).
46. B. S. Bollepalli, M. Khan and F. Cerrina, 'Image formation in extreme ultraviolet lithography and numerical aperture effects,' *J. Vac. Sci. Technol. B*, vol. **17**, no. 6, p. 2992, (1999).

15

Mask-substrate alignment via interferometric moiré fringes

E. E. MOON, Massachusetts Institute of Technology, USA

DOI: 10.1533/9780857098757.466

Abstract: Highly robust alignment is achieved with formation and analysis of interferometric moiré fringes. High signal-to-noise is attained by illuminating and imaging at an oblique angle, avoiding contribution of reflections to an image. Detectivity is below the nanometer level, due to displacement magnification in the moiré pattern and spatial-phase analysis. The moiré fringe period is independent of illumination wavelength. Alignment marks can be placed on the wafer backside and imaged with infrared illumination, providing a permanent, unambiguous reference for multi-level alignment.

Key words: interference, fringe, moiré, alignment, metrology.

15.1 Introduction

Alignment, or control of spatial relations between multiple objects, is essential in many aspects of nano-fabrication. Manufacture of semiconductor processors, for example, would not be possible without alignment of dozens of masks, often with tolerances that are a small fraction of the minimum feature width. Numerous alignment methods, at least in principle, are able to measure sub-10 nm mask–substrate displacements. However, as device dimensions shrink into the sub-10 nm regime, there is a need for a robust alignment technology capable of detecting displacement on the nanometer scale, or smaller.

In typical manufacturing conditions, performance of alignment methods is degraded by process-induced perturbations of the optical properties of substrate alignment marks. Variations originate from uneven resist flow over the marks, asymmetry, or unequal line–space ratios induced by etching and deposition, and non-homogeneous refractive indices. In practice, such disturbances usually reduce alignment detectivity by an order of magnitude. A common factor in such methods is encoding alignment information in the amplitude of the interferometric signals, which is susceptible to noise and perturbations and undermines the promise of high sensitivity inherent to interferometry.

Section 15.2 describes several representative methods that utilize optical interference, in combination with diffracting marks on a substrate surface. In Sections 15.3–15.7 we introduce an alignment method, called interferometric-spatial-phase imaging (ISPI), which encodes position information in the frequency domain, at a predetermined *spatial frequency*. This approach enables filtering of spurious, disruptive spatial frequencies originating from process-induced and environmental disturbances. At the same time, ISPI provides sub-nanometer detectivity, high signal-to-noise ratio, and large dynamic range, as well as immunity to numerous other deleterious conditions. Section 15.7 illustrates a variation of ISPI that uses alignment marks on the backside of a substrate.

15.2 Background to alignment methods

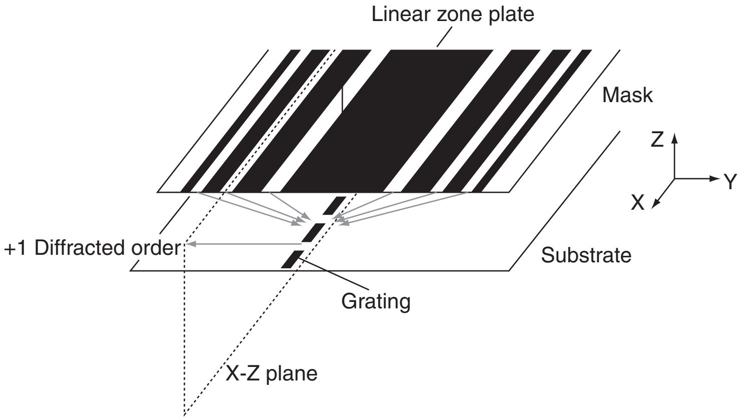
The simplest form of alignment uses optical microscopes to observe geometric marks, such as crosses on a mask, within a symmetric structure on a substrate (e.g., boxes). Typically, alignment is limited to a few hundred nanometers, but with signal processing and averaging over many such marks, alignment detectivity can be improved to ~50 nm.[1] In the following sections, we review methods to extend alignment beyond the capability of such box-and-cross methods.

15.2.1 Linear zone-plates

A linear zone plate (LZP) is a type of planar lens, since it causes an incident plane wave to be focused to a narrow stripe. This stripe can be scanned over a small mark, producing a position-dependent variation in the intensity of a return signal.

In one class of alignment methods, LZPs (Fig. 15.1) were placed on a mask above a diffracting mark, consisting of a narrow grating on a substrate. A first-order diffracted beam from the grating provided the alignment signal, which was detected by a photodiode. The intensity of the diffracted beam was measured during a lateral scan of a substrate stage, where the shape of the signal during the scan was a convolution of the widths of the grating and the focal spot. The aligned condition was taken to be the centroid of the peak.[2] Variations on this scheme allowed faster data acquisition rates by scanning the laser illumination between two sides of the mark,[3–5] instead of mechanically scanning the stage.

A refinement of this method used a pair of adjacent substrate gratings incorporating a π-phase shift.[6] Here the observed signal consisted of two peaks with a central minimum, due to cancellation of light between the two π-phase-shifted gratings. The amplitude minimum was interpreted to



15.1 Alignment scheme by Fay *et al*. using a LZP on a mask and a narrow, orthogonal diffraction grating on a substrate. A normal-incidence laser beam illuminated the marks while the zone plate was scanned over the substrate grating. The alignment signal was the intensity of a first-order diffraction from the substrate mark, diffracted out of the depicted plane.

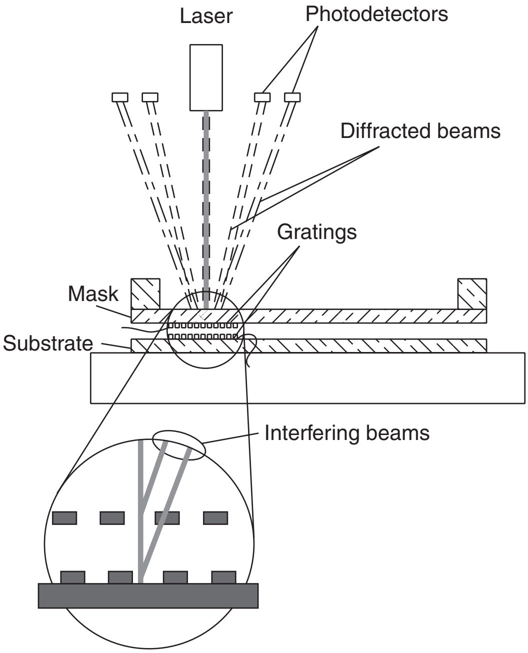
be the aligned condition. In this case, a mirror scanned the illumination over a small angle to obtain a time-dependent signal, and the substrate was scanned to find the aligned position. Repeatability was found to be high, with $3\sigma < 1$ nm, when using no resist or other process coatings.

A difficulty with zone plate methods was limited focal depth, which implied unintended signal variations outside a small range of gaps. Since the alignment signal depended only upon amplitude, in the absence of supplementary information, misalignment was indistinguishable from changes in gap. Other constraints included the alignment optics blocking an exposing beam and the ability to measure only a point of symmetry, instead of arbitrary misalignments.

15.2.2 Interferometers

A large number of methods were developed in industry and academia to use interferometry with diffracting marks on a mask and a substrate. Here we will briefly examine two examples:

In one method (Fig. 15.2), a linear grating was written on a mask and another on a substrate, with the same period in each mark.[7] A laser illuminated the gratings at normal incidence, resulting in diffracted +1 and −1 orders. Each grating produced a set of orders diffracted at the same angle, resulting in interference between like orders. Photodetectors measured the intensity of interference at symmetric locations above the mask. The



15.2 Alignment scheme of Flanders, Smith, and Austin. Similar backdiffracted orders (e.g., +1 from the mask and +1 from the substrate) interfere. The interferometric intensity was measured by photodetectors. Alignment was indicated by the voltage difference between two photodetectors measuring symmetric (+1 and −1) diffracted order groups.

amplitude of the detected signals from symmetric orders was compared, and related to the relative position of the gratings. In this way, a differential signal would indicate alignment: when one grating was directly on top of the other grating, there would be a minimum differential signal; otherwise, the differential signal would vary sinusoidally with misalignment. However, since the amplitude at each photodetector resulted from the path length difference of two diffracted orders, one from the upper grating, and one from the lower grating, the detected amplitude could result from either a lateral or a vertical motion of a grating. As with the zone plate-based schemes, the mask–substrate gap was coupled with the alignment signal, and any asymmetric processing or blazing of a grating would affect the alignment signal.[8] Another difficulty was matching the response of multiple photodetectors.

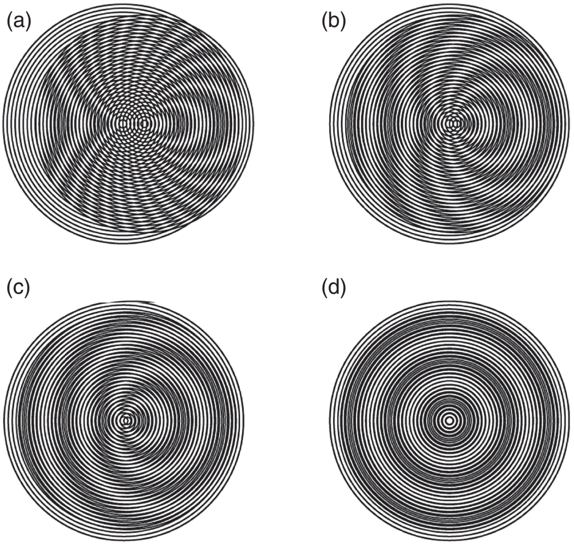
In a variation of this scheme, the intensity resulting from interference of multiple diffracted beams was measured by a single photodetector, positioned at a small angle from the surface normal.[9] During a scan of the substrate, the detected amplitude underwent a sinusoidal variation from interference of similar diffracted orders, with an envelope determined by the overlap of the two marks. In this manner, the centroid of the signal was calculated and used as a coarse positioning estimate, while fine position was determined by the amplitude within one cycle near the centroid of the envelope. A drawback of this method was the necessity to scan the wafer position to generate the alignment signal, and subsequently rely upon the stage interferometer and mechanics to return the stage to the intended position. A separate interferometer attached to the stage could resolve position to ~1 nm, but the substrate position was still subject to thermal drift and other disturbances, which could not be observed with this interferometer. As in other amplitude-sensing interferometric schemes, blazing or other asymmetric changes to the substrate grating would perturb the alignment signal.

15.2.3 Moiré patterns

Moiré patterns provide highly sensitive indications of position and other characteristics of underlying structures. In general, moiré is produced by the superposition of two or more periodic structures, resulting in spatial frequencies at the difference of the initial frequencies: similar initial frequencies produce a small frequency difference and consequently a large fringe period. Since the initial structures are periodic, a shift in position by a structure through one period results in a phase condition, which applies equally to the physical structures and the resulting fringes. Enhanced position sensitivity arises from the synchronous motion of the large-period moiré and the grating, resulting in a greater relative distance traversed by the moiré.

The first known scientific application of moiré was in 1874, when Lord Raleigh found that the superposition of two gratings produced a set of broad stripes, or low-frequency fringes, which was indicative of the relative quality of the gratings.[10,11] Despite the sensitivity of moiré, very little use was made of it in displacement measurement until the 1950s, when Guild[12,13] applied moiré to machine control, by devising moiré-based optical encoders to measure linear translation of stages. The principle of moiré encoding has widespread use in contemporary linear encoders, although grating quality presently limits their accuracy to a few tens of nanometers. A significant advance in ruling engines, the novel scanning beam interference lithography (SBIL), or 'NanoRuler', writes a grating with an interferometric 'brush,' achieving high-speed writing and nanometer placement accuracy over a 300 mm diameter substrate.[14,15]

The first use of moiré for mask–substrate alignment was by King and Berry in the early 1970s.[16] In the King and Berry apparatus, moiré fringes were produced between circular gratings, as depicted in Fig. 15.3, and were

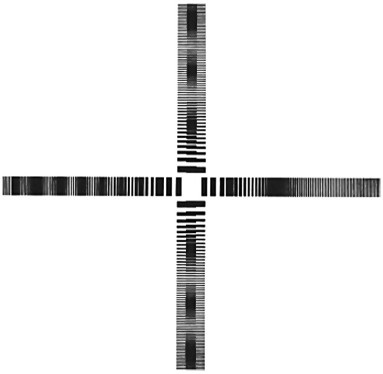


15.3 Schematic of circular moiré patterns illustrating the method used by King and Berry, as well as Goodberlet.[17] The grating periods are 4.0 and 4.4 μm. The moiré fringe period is (4.4 × 4.0)/(4.4–4.0) = 44 μm. (a) Displacement in the X-direction by three grating periods (b) Displacement in X by one grating period (c) Displacement in X by 0.5 grating period (d) Zero displacement. Circular moiré fringes effectively indicate the direction of offset (along the vector that symmetrically bisects the fringe pattern), as well as the magnitude of displacement. Displacement is proportional to the number of fringes that intersect the vector between the centers of each grating.

observed with a 20× normal-incidence, brightfield, white-light microscope. Given that translation through a fringe period was 44 μm, which corresponded to a translation of the circular grating by the average of the two grating periods, one would expect alignment accuracy by eye to be limited to about 1/20 of a fringe cycle, or 4.2 μm/20 = 0.21 μm. This number is consistent with their demonstrated alignment of <1 μm.

Another implementation of moiré was in the shape of a cross (Lyszczarz *et al.*, 1981)[18] as a coarse alignment method in conjunction with an interferometric alignment method. The moiré was produced by interference of monochromatic light from orthogonal LZPs on both the mask and substrate, Fig. 15.4. Alignment sensitivity was <250 nm.

In the 1990s, Moel *et al.* used rectilinear moiré fringes compared to fiducial marks to facilitate automated analysis and e-beam generation of the moiré alignment marks.[19] In this scheme, called ‘on-axis interferometric (OAI)’ alignment, a grating of period p_1 on a mask was superposed upon a grating of similar period p_2 on a substrate. To determine alignment, the



15.4 Schematic of a moiré cross produced by a LZP arrangement. LZPs were on both mask and substrate. The illustration shows alignment in the Y-direction and misalignment in the X-direction.

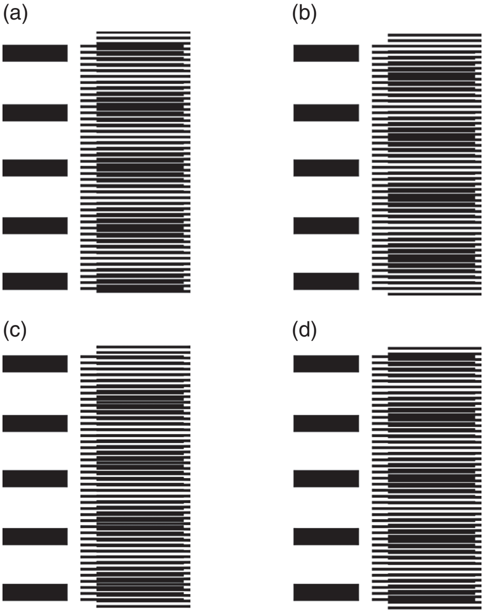
spatial-phase difference was found between a set of moiré fringes, detected by a CCD or CMOS camera. The fringes result from overlay of p_1 and p_2 gratings, and a coarse fiducial mark etched into the substrate, which was designed to have the same spatial period as the fringes, as shown in Fig. 15.5. Additional sensitivity was obtained by analyzing the moiré fringes in the frequency domain. Alignment detectivity was < 10 nm.

This was the first demonstration of frequency domain analysis for alignment purposes, and is also used in the ISPI system, as described in the following section.

15.3 Fundamentals of interferometric-spatial-phase imaging (ISPI)

ISPI combines the strengths of darkfield optical microscopy with interferometric moiré fringes, formed between specialized diffraction marks on a mask and substrate. As shown in Fig. 15.6, marks on the mask consist of a set of gratings with period p_1. A complementary checkerboard mark is on the substrate, containing a period p_2 (and an orthogonal period p_h). Illumination is provided by multiple collimated wavelengths, conveyed on a fiber optic, and reflected from a mirror towards the marks. Light is returned to a microscope at oblique angles from the p_1 and p_2 marks. Diffracted beams from the two marks result in interferometric moiré fringes with period P_F, imaged onto a CCD camera. The low-spatial-frequency moiré fringes are analyzed using a robust frequency domain algorithm that filters all frequency components except for the known fringe frequency.

A characteristic feature of ISPI is the use of an oblique angle to selectively retrieve signals originating only from specific diffracted beams. Most

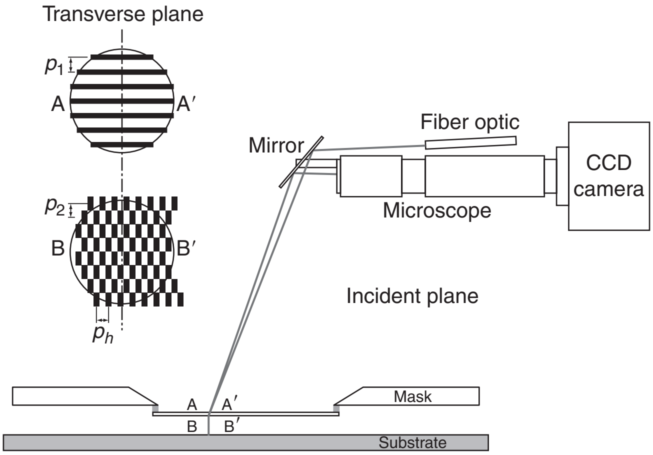


15.5 Illustration of alignment marks used in the OAI alignment scheme. The substrate contains a set of fiducial marks and a grating. The mask contains only a grating. Displacements along the *k*-vector of the gratings are indicated in the sequence (a)–(d), corresponding to a shift of one period of the mask grating. Alignment was signified by the relative position of the set of moiré fringes, compared to a set of fiducial marks on the substrate. (a) Mask grating displaced by one grating period in +Y-direction. (b) Displacement of half a grating period. (c) Displacement of a quarter grating period. (d) Aligned position. Note that displacement by one period realigns the moiré fringes to the fiducial marks. This ambiguity was resolved with conventional cross and box marks. Process layers, such as resist, affected the fiducial marks differently than the moiré fringes, inducing alignment errors under common conditions.

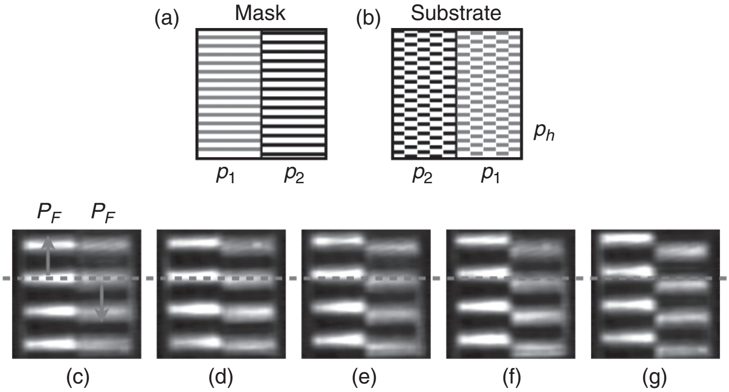
importantly, this viewing angle ensures that the CCD camera does not collect light reflected from the mask and substrate. As the magnitude of the reflected beam is much stronger than any diffracted beams, and carries little useful information, its removal results in a significant increase in fringe contrast and signal-to-noise ratio.

The alignment marks are designed to produce moiré fringes that are independent of illumination wavelength, refractive index, and mask–substrate gap. These features are discussed in detail in Section 15.3.

Another key feature of ISPI is encoding position in moiré fringes. Moiré provides high magnification of displacement, and is a critical component in high detectivity.

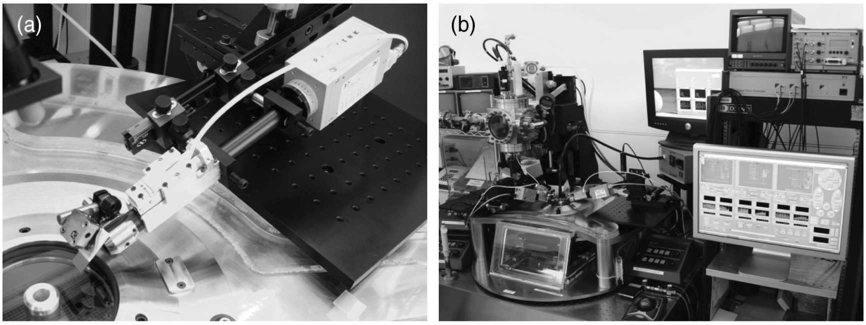

Transverse plane
p_1
A
A′
p_2
B
B′
p_h
Mirror
Fiber optic
CCD camera
Microscope
Incident plane
A
A′
B
B′
Mask
Substrate

15.6 Schematic of ISPI scheme. Marks on a mask and a substrate diffract incident light to a microscope, where moiré fringes are imaged onto a CCD camera. Reflected light does not contribute to fringe formation, resulting in high signal-to-noise ratio. Moiré fringes magnify motion of the mask or substrate.


(a)
Mask
p_1
p_2
(b)
Substrate
p_h
p_2
p_1
P_F
P_F
(c)
(d)
(e)
(f)
(g)

15.7 ISPI alignment information is contained in the spatial-phase discontinuity of a pair of interference patterns, produced by (a) $\{p_1, p_2\}$ grating sets on the mask and (b) $\{p_2, p_1\}$ marks on a substrate. Images (c)–(g) illustrate counter-moving fringes with spatial-phase disparities corresponding to displacements of 50 nm.

A further unique aspect is a *phase reference* produced by moiré fringes. As shown in Fig. 15.7, a full ISPI mark consists of two parts: one that has a p_1 grating region superimposed on a p_2 region, and an adjacent part that superimposes p_2 over p_1 regions. The resulting sets of moiré fringes have an identical spatial frequency P_F, but move in opposite directions in response to physical displacements. In this manner, both sets of fringes are subject to



15.8 (a) Photograph of an ISPI microscope. (b) Photograph of three ISPI microscopes obliquely viewing ISPI marks on a mask. Illumination is conveyed through single-mode fiber optics, collimated, and reflected by a metalized silicon mirror to the mask and substrate at an 18° angle. Diffracted beams return at a 22° angle, and are imaged onto CCD cameras by microscopes at a working distance of 110 mm from the mask.

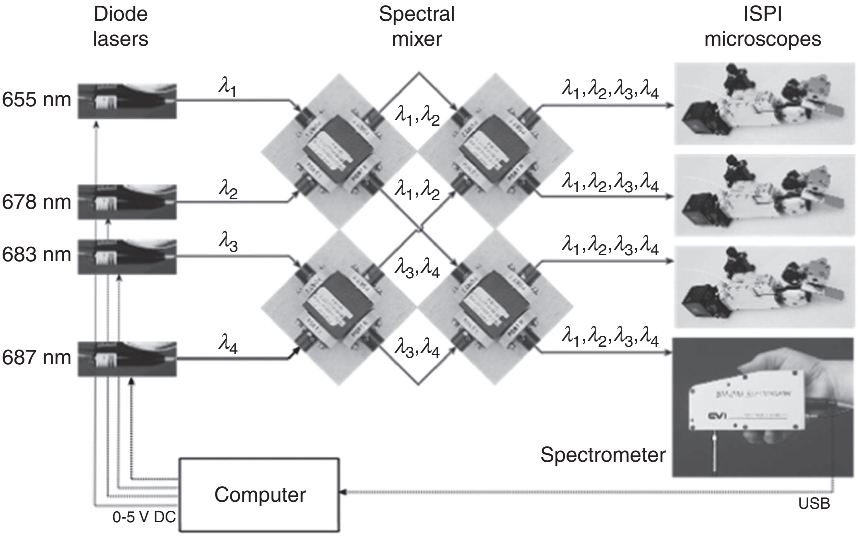
the same optical variations (e.g., resist layer or changes in refractive index in the optical path), resulting in cancellation of many perturbations.

15.4 Implementation of moiré

In this section we will describe one implementation of Interferometric Spatial-Phase Imaging and its detectivity. A benefit of ISPI is that the optics can be relatively simple and low cost, the image analysis is not exceptionally complicated, yet the resulting detectivity is in the sub-nanometer regime.

15.4.1 Imaging system

Low-magnification, low-NA microscopes (Fig. 15.8) are key components of the ISPI system. Low-pass spatial filtering of the diffracted beams by low-NA optics eliminates many small process-induced defects from the image. Other advantages of low-NA microscopes are the consequent long working distance, which allows the microscopes to be sufficiently far away from the mask to avoid obstructing any exposing radiation, and the large focal depth, which allows measurements to be taken over a large range of gaps. The microscopes are based on commercial machine vision lenses, with modifications to the illumination input. The microscope working distance is 110 mm, the focal depth is 150 μm, and resolution is 11 μm. Light is incident at a 4° angle from above the optical axis using a custom designed fiber holder. Imaging and illumination paths are deflected towards the mask and substrate with a silicon mirror. The mirror angles are adjusted with a kinematic tip–tilt stage, microrail components, and a custom rail attachment. The fiber optic is a single-mode design, with a 3 μm core and a pigtail collimator. The camera uses a 6.6 megapixel CMOS sensor.


Diode lasers
Spectral mixer
ISPI microscopes
655 nm
678 nm
683 nm
687 nm
λ_1
λ_2
λ_3
λ_4
λ_1,λ_2
λ_1,λ_2
λ_3,λ_4
λ_3,λ_4
$\lambda_1,\lambda_2,\lambda_3,\lambda_4$
$\lambda_1,\lambda_2,\lambda_3,\lambda_4$
$\lambda_1,\lambda_2,\lambda_3,\lambda_4$
$\lambda_1,\lambda_2,\lambda_3,\lambda_4$
Spectrometer
Computer
0-5 V DC
USB

15.9 Schematic of the multi-line, simulated broadband light source. (Images of components are to different scales.) A set of four fiber-coupled beamsplitters mixes the four input wavelengths equally in four outputs. ISPI microscopes use three outputs, and the fourth is used for spectral analysis.

15.4.2 Broadband illumination

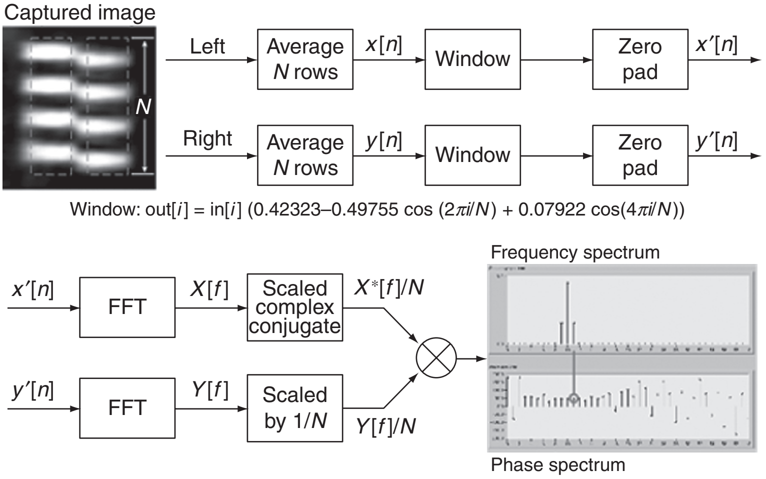
To obtain high-contrast moiré fringes it is desirable to use spatially collimated light. However, single laser wavelengths present the problem of thin-film interference, which can result in extinction of the signal at certain film thicknesses. To avoid such issues, we use a multi-line laser apparatus, closely spaced in wavelength – which simulates broadband illumination while retaining high spatial collimation. As we will discuss further in Section 15.5.1, ISPI fringe formation is wavelength-independent, allowing multiple wavelengths to produce ISPI fringe sets that superimpose with the same spatial phase. Although fringes at one wavelength may undergo variations in amplitude from thin-film interference, the multiple-wavelength fringe sets provide a consistently high-contrast fringe signal. The implementation of the multi-line illumination source is shown in Fig. 15.9.

The multi-line light source consists of four diode lasers (655, 678, 683, and 687 nm, 30 mW a piece) whose spectrum is mixed in a combination of four beam splitters. The beam splitters distribute equal amounts of each spectral line to the four fiber outputs. The ISPI microscopes use three of these outputs, and the fourth output is used by a compact spectrometer, whose output is transferred to a computer and used to ascertain the wavelength and power transmitted through the spectral mixer. The power of each laser diode is subsequently adjusted to obtain a desired spectral content.

15.4.3 Image analysis

One of the primary advantages of encoding alignment at a predetermined spatial frequency is that contributions from spurious spatial frequencies can be filtered, while retaining the desired position information. Defects in the alignment marks may contain spatial frequencies that are not removed by low-pass optical spatial filtering. However, such defects are very rarely periodic at the same frequency as the fringes. Since analysis of the spatial-phase difference is at a single, known spatial frequency, contributions of spurious spatial frequencies can be excluded. As a result, we extract from the image only spatial frequencies that correspond to the predetermined moiré frequency.

A block diagram of the algorithm is shown in Fig. 15.10. Within each captured image, two regions of interest (ROI) are extracted, with the length of each region set equal to N, the number of pixels perpendicular to the fringes. Note that the number of fringes within the sample length N can be a non-integer number. The width of each region is constrained by the boundary between the two fringe sets. In left and right ROI, the pixels are averaged along every row, forming one-dimensional signals, $x[n]$ and $y[n]$. Each one-dimensional signal is multiplied by a Blackman–Harris windowing function and zero padded. The FFT produces $X[f]$ and $Y[f]$, from which


Captured image
N
Left
Average N rows
$x[n]$
Window
Zero pad
$x'[n]$
Right
Average N rows
$y[n]$
Window
Zero pad
$y'[n]$
$x'[n]$
FFT
$X[f]$
Scaled complex conjugate
$X^*[f]/N$
$y'[n]$
FFT
$Y[f]$
Scaled by 1/N
$Y[f]/N$
Frequency spectrum
Phase spectrum

Window: $\text{out}[i] = \text{in}[i]\,(0.42323 - 0.49755\cos(2\pi i/N) + 0.07922\cos(4\pi i/N))$

15.10 Block diagram of the basic phase analysis algorithm used to determine the phase difference between two sets of ISPI fringes.

we find $X^*[f]/N$ and $Y[f]/N$. The product of $X^*[f]/N$ and $Y[f]/N$ yields exponentials with the difference of the phase at every frequency, seen in plots of the magnitude and phase in Fig. 15.5, where the phase difference of the two fringe sets is given by the value of the phase spectrum at the index corresponding to the frequency peak. Sub-pixel sensitivity is obtained with this phase analysis algorithm.

The phase value is converted into a displacement with the expression

$$\Delta x = \frac{\phi P_{\text{av}}}{8\pi} \qquad [15.1]$$

where period P_{av} is given by

$$P_{\text{av}} = \frac{2p_1 p_2}{p_1 + p_2} \qquad [15.2]$$

and p_1 and p_2 are the grating periods.

Note that according to Equation [15.1], any phase disparity is measured with equal accuracy.

Including the effects of noise, the phase can be measured with an efficient estimator, such as the discrete Fourier transform used in the fast Fourier transform. The limit on phase estimation is given by the Cramer–Rao bound, which determines the minimum variance of a physical variable in the presence of noise, indicating that the standard deviation of the phase estimate is

$$\sigma_\phi = \sqrt{\frac{2(N-1)}{(\text{SNR})(N)(N+1)}} \approx \sqrt{\frac{2}{(\text{SNR})(N)}} \qquad [15.3]$$

for large N, where N is the number of samples, and SNR is the signal-to-noise ratio. The SNR is given by

$$\text{SNR} = \frac{A^2}{\sigma^2} \qquad [15.4]$$

where A is the amplitude of the ISPI fringes and σ is standard deviation of the noise.

If the standard deviation of the phase is given by σ_φ, then the standard deviation of the position measurement is given by

$$\sigma_x = \frac{\sigma_\phi P_{\text{av}}}{8\pi} \qquad [15.5]$$

In ISPI we obtain high SNR: the dynamic range of a simple 8-bit camera is 0 to 255, and the intensity of the lasers is adjusted so that the amplitude of the ISPI fringes essentially spans the entire dynamic range. Noise is low, with $\sigma \sim 5$, yielding SNR = 2500. The number of pixels that might be used along one set of fringes with a simple (640 × 480)-pixel camera is $N = 300$. The standard deviation of the phase estimate is

$$\sigma_{\phi}=\sqrt{\frac{2(300-1)}{(2500)(300)(300+1)}}=0.001628 \qquad [15.6]$$

Therefore the standard deviation in the position measurement with a basic (640 × 480)-pixel-resolution, 8-bit camera can be expressed

$$\sigma_x=\frac{(0.001628)(1)}{8\pi}=0.065\text{ nm} \qquad [15.7]$$

Improved position resolution is achieved in a straightforward manner using megapixel cameras with higher bit depth.

15.5 Characteristics of moiré fringe formation

Here we illustrate several key characteristics of the Interferometric Spatial-Phase Imaging method. We show how these features originate from simple, but very specific, combinations of diffracted beams.

15.5.1 Wavelength independence

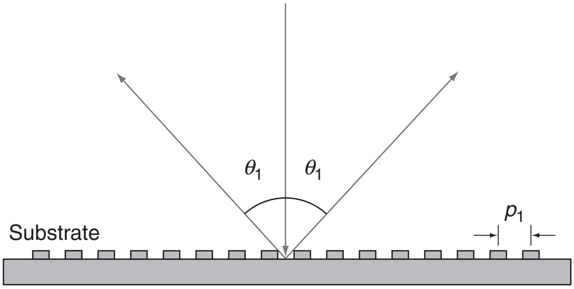
In this section we will derive the moiré fringe formation from basic diffraction equations and note the cancellation of wavelength dependence due to multiple diffractions from mask and substrate marks.

Consider diffraction of an incident beam from a periodic structure, as illustrated in Fig. 15.11. In the transverse plane, diffraction will occur from the substrate period p_1 at angles given by

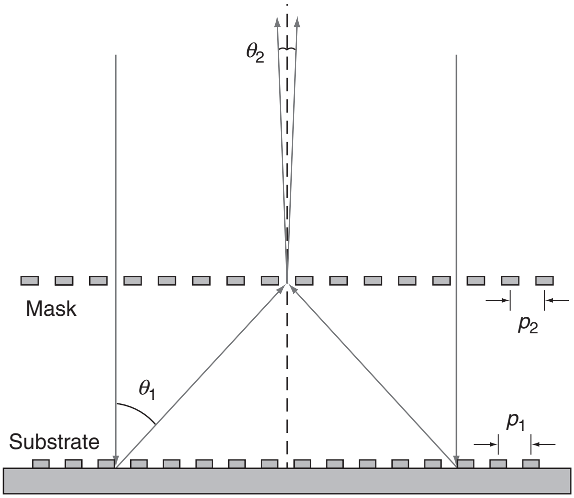
$$\theta_1=\sin^{-1}\left(\frac{m\lambda}{p_1}\right) \qquad [15.8]$$

where λ is the wavelength of illumination and m is the order of diffraction.

The beams that diffract from the substrate grating will diffract again at the mask grating, as illustrated in Fig. 15.12.



15.11 Schematic of diffraction from a substrate grating in the transverse plane.



15.12 Schematic of double diffraction from two different grating periods, resulting in interference moiré above the mask plane.

The angle of diffraction from the mask grating p_2 is given by

$$\theta_2 = \sin^{-1}\left(\frac{m\lambda}{p_2} + \sin\theta_1\right) \qquad [15.9]$$

where θ_1 is the angle of the beam diffracted from the period p_1 on the substrate, and θ_2 is the angle of the beam diffracted by the mask period p_2.

As an example, let $p_1 = 1.000$ µm, $p_2 = 1.025$ µm, $\lambda = 690$ nm, and $m = +1$,

$$\theta_1 = \sin^{-1}\left(\frac{(+1)(0.690)}{1.000}\right) = 43.63°$$
$$\theta_2 = \sin^{-1}\left(\frac{(-1)(0.690)}{1.025} + \sin 43.63\right) = 0.96° \qquad [15.10]$$

It is apparent that the double diffraction causes two large-angle diffractions to result in a small final diffraction angle.

A second doubly-diffracted beam will also emerge that is symmetric with the first, as a result of reversal of the sign of the first orders.

$$\theta_1 = \sin^{-1}\left(\frac{(-1)(0.690)}{1.000}\right) = -43.63°$$
$$\theta_2 = \sin^{-1}\left(\frac{(+1)(0.690)}{1.025} + \sin 43.63\right) = -0.96° \qquad [15.11]$$

Thus, two components of the source illumination diffract from a mark on the substrate, meet at the mask, and diffract again. Each beam that contributes to image formation is diffracted in first order twice. A beam diffracts in first order from p_1 and then from p_2, or first from p_2 and then p_1. The same fringe period is produced regardless of the sequence of the +1 and −1 diffractions, resulting in identical periods in the interference pattern. The interference of these beams, diffracting at small angles, forms a fringe pattern with period P_F, given by

$$P_F = \frac{\lambda}{2\sin\theta_2} \qquad [15.12]$$

where the θ_2 is the final diffraction angle. Expressing θ_2 from Equation [15.4] in terms of p_1 and p_2

$$\theta_2 = \sin^{-1}\left(\frac{m\lambda}{p_2} + \sin\theta_1\right) = \sin^{-1}\left(\frac{m\lambda}{p_2} + \frac{-m\lambda}{p_1}\right) \qquad [15.13]$$

or, if we assume only the first-order $|m| = 1$

$$\theta_2 = \sin^{-1}\left(\frac{\lambda}{p_2} - \frac{\lambda}{p_1}\right) = \sin^{-1}\lambda\left(\frac{1}{p_2} - \frac{1}{p_1}\right) = \sin^{-1}\lambda\left(\frac{p_1 - p_2}{p_1 p_2}\right) \qquad [15.14]$$

Substituting θ_2 into Equation [15.5] yields

$$P_F = \frac{1}{2(p_1 - p_2/p_1 p_2)} = \frac{p_1 p_2}{2(p_1 - p_2)} \qquad [15.15]$$

For arbitrary values of p_1 and p_2 the moiré fringe period is

$$P_F = \frac{p_1 p_2}{2|p_1 - p_2|} \qquad [15.16]$$

For $p_1 = 1.00$ μm and $p_2 = 1.025$ μm this equation predicts a fringe period $P_F = 20.5$ μm. This expression for the fringe period is half the fringe period for geometric moiré, due to the removal of the zero-order (the reason for removal of the zero-order will be explained in Section 15.5.5 on Littrow-angle viewing). Note that the contributions from the wavelength of illumination cancel, and the expression for P_F is achromatic.

Multiple diffraction paths and multiple illumination wavelengths lead to formation of several sets of moiré fringes, all of which superimpose with the same spatial phase.

Moiré displacement magnification is given by the expression

$$M = \frac{O_m C_p P_F}{P_{\text{av}}} \qquad [15.17]$$

where O_m is the optical magnification, C_p is another factor of two from the counter-propagating fringes, P_F is the fringe period, and P_{av} is the average grating period. The average grating period P_{av} is given by

$$P_{\text{av}} = \frac{2 p_1 p_2}{p_1 + p_2} \qquad [15.18]$$

where p_1 and p_2 are the grating periods, and are typically on the order of 1 μm. Usually, the moiré magnification is one or two orders of magnitude larger than the grating periods. For instance, if grating periods $p_1 = 2.000$ and $p_2 = 2.050$ μm, producing fringe period $P_F = 41.000$ μm, or grating periods $p_1 = 1.000$ and $p_2 = 1.025$ μm, producing fringe period $P_f = 20.500$ μm. The mark length is typically 400 μm, resulting in 400 μm/41 μm = 9.75 fringes or 400 μm/20.5 μm = 19.5 fringes along the mark. For the ~2 μm gratings, the fringe magnification is $M = 972$, indicating that a displacement of the mask by 10 nm causes a fringe displacement of 9.72 μm on the CCD sensor (one CCD pixel is 3.5 μm).

As we have seen, one of the desirable aspects of the design is that within the transverse plane, light that contributes to the interference pattern passes through the gratings by diffracting twice: once from the top grating and once from the bottom grating, or, once from the bottom grating and once from the top grating. The order does not matter and the interference fringes resulting from both diffraction paths super impose, with no phase difference. The result of this double diffraction is that the angle at which light diffracts from one grating is compensated by the angle of diffraction from the second grating. This relation leads to a critical point: within a range, the same fringe pattern is obtained with a wide spectral bandwidth of light: any wavelength of illumination produces an identical fringe pattern. Consequently, one may use the same alignment method with many visible wavelengths, or infrared illumination, as described in Section 15.7.

15.5.2 Refractive-index independence

In this section we will derive the phase invariance of a layer of resist on top of the diffraction marks on the substrate. We then show experimental verification of the immunity of the ISPI signal to layers of resist.

Consider a layer of resist covering an alignment mark on the substrate, as illustrated in Fig. 15.13.

The equations that describe the effect are

$$n_a \lambda_a = n_r \lambda_r \tag{15.19}$$

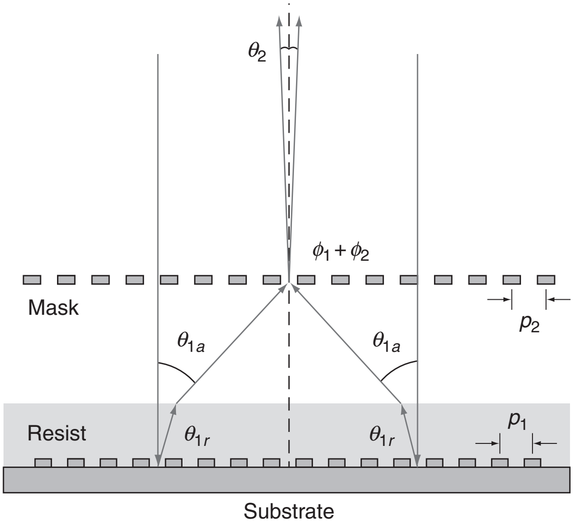
where the subscript '*a*' indicates air and the subscript '*r*' indicates resist, and

$$\theta_{1r} = \sin^{-1}\left(\frac{n_a \lambda_a}{n_r p}\right) \tag{15.20}$$

where θ_{1r} is the angle of the first-order diffracted beam inside the resist. By Snell's law

$$n_a \sin\theta_{1a} = n_r \sin\theta_{1r} \tag{15.21}$$

Solving for θ_{1a},



15.13 Sketch of diffracted beams inside resist, taking note of the refraction of the diffracted beams. The wavelength change within the resist introduces a compensating factor, which allows cancellation of the index of refraction of resist in the expression for the refraction angle. The angle of a first-order diffracted beam above the resist is the same, with or without resist, although the optical path length changes in the presence of resist. The fringe frequency is unaffected by the resist because fringe formation depends only on the angles between the diffracted beams. The fringe phase is unaffected as a result of equal and opposite phase changes in the two paths.

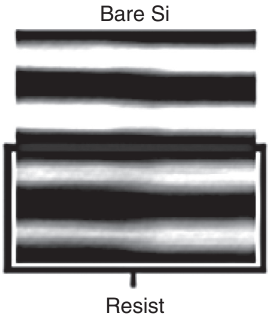
$$\theta_{1a} = \sin^{-1}\left(\frac{n_r}{n_a}\sin\theta_{1r}\right) \qquad [15.22]$$

Upon substitution for $\sin\theta_{1r}$,

$$\theta_{1a} = \sin^{-1}\left(\frac{n_r}{n_a}\left(\frac{n_a\lambda_a}{n_r p}\right)\right) \qquad [15.23]$$

or

$$\theta_{1a} = \sin^{-1}\left(\left(\frac{\lambda_a}{p}\right)\right) \qquad [15.24]$$



15.14 ISPI image of fringes from an alignment mark etched into a substrate. Half of the mark was covered with an overlayer. The other half was left bare. Experiments were done in this manner with overlayers of resist, as well as polysilicon, and aluminum.

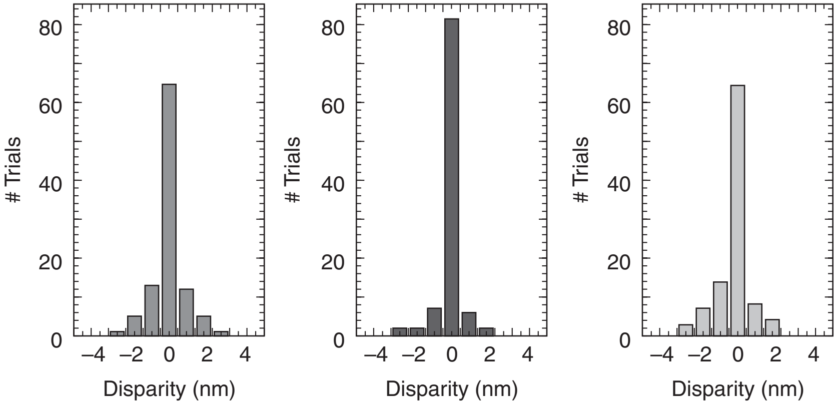
which is equivalent to θ_1, the diffraction angle that would have occurred without any resist. Although the resist layer will cause a displacement of a diffracted beam, the angle of diffraction (hence the moiré fringe frequency) will remain the same, with or without the resist. Since the phase is changed by equal and opposite amounts in the beam paths within a constant-thickness resist, the alignment signal is unaffected by a resist layer.

The phase invariance of the ISPI signal is supported in an experiment: an ISPI alignment mark was etched into silicon, and half of the mark was covered with an overlayer (Fig. 15.14). The mark was observed with an ISPI microscope, and the resulting fringe patterns for the upper and lower halves of the mark were analyzed separately, using the ISPI phase algorithm.

The effects on alignment of three types of overlayers – photoresist, polysilicon, and aluminum – were investigated. Substrate alignment marks consisted of gratings $p_1 = 3.7$ µm and $p_2 = 4.0$ µm, etched 100 nm deep into a Si substrate. The marks, which occupied an area 200 × 200 µm, were covered with a given overlayer, and the overlayer was subsequently removed from one half of the substrate alignment mark (by exposure of coarse features aligned to the mark, and subsequent etching of the polysilicon and aluminum, or, for resist, development and removal). Overlayers of resist and polysilicon decreased fringe contrast, while aluminum increased contrast.

Figure 15.15 displays the results of measurements of the difference in the spatial-phase discontinuity between the half of the alignment mark with the overlayer and the half without the overlayer.

The 100 measurements shown in Fig. 15.10 for each overlayer yielded a mean offset, δ, as well as a standard deviation, σ. The measured effects of the overlayers were $\delta = 0.1$ nm, $\sigma = 2.4$ nm for 1 µm of photoresist, $\delta = 0.1$ nm, $\sigma = 1.2$ nm for 300 nm of deposited polysilicon, and $\delta = 0.2$ nm,



Resist:
$\delta = 0.1$ nm
$\sigma = 2.4$ nm

Polysilicon:
$\delta = 0.1$ nm
$\sigma = 1.2$ nm

Aluminum:
$\delta = 0.2$ nm
$\sigma = 2.6$ nm

15.15 Immunity of ISPI to overlayers of resist, polysilicon, and aluminum. Layers of resist, polysilicon, and aluminum are shown to cause minimal degradation of the alignment signal. Mean alignment errors are less than the detectivity of the alignment measurements.

$\sigma = 2.6$ nm for 20 nm of deposited aluminum. Note that the ISPI phase measurement is invariant with the thickness of resist, since the diffraction angle is independent of the layer thickness. Since the two halves of the ISPI mark are in close proximity (200 μm or less) the only variation of some concern to ISPI would be a ramp in the thickness of resist across the mark. Resist is typically flat and consistent in thickness to the nanometer level over the dimensions of the mark, and this is not observed to be a problem in practice.

The significance of the results is that an overlayer does not appear to affect the *spatial phase* of the interference pattern, although it does affect the image contrast. Since we signify alignment by the matching of the spatial phase between two sets of interference fringes, the effect of overlayers on alignment is shown to be negligible.

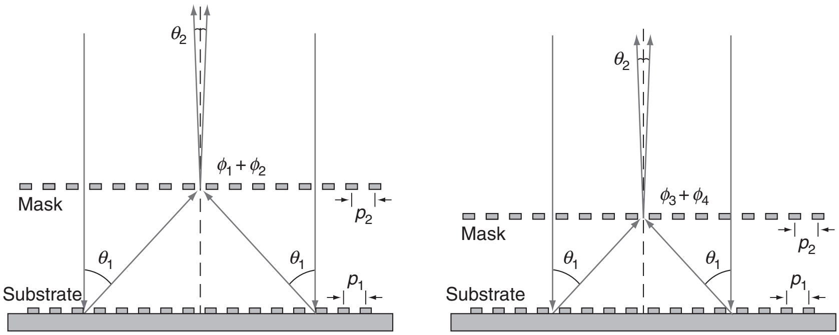
We have demonstrated that any overlayer that causes a phase change on one side of the mark, will also, in all likelihood, affect the phase on the other side of the mark in exactly the same way since each set of fringes is formed by the interference of symmetric diffractions from mask and substrate gratings. In distinction, in an amplitude-based alignment method, the alignment signal could depend on a plethora of factors such as the gap, overlayer type and thickness, transmission of the mask, output stability of the laser, etc. The only situation in which the phase symmetry in ISPI would be broken is for extreme variations in the layer thickness occurring over the small extent of the alignment mark. The data suggest that thickness variations of a magnitude sufficient to distort the signal are not encountered in practice.

However, it should be noted that processes that induce a blazing of the gratings, i.e., an asymmetric covering of the gratings, will shift the grating lines in one direction. This is equivalent to moving the substrate gratings, which we know will cause the observed fringes to shift in opposite directions. Thus, in the special case of asymmetric overlayers, we expect a spatial offset. However, the ability of ISPI to detect arbitrary offsets with equal accuracy should allow one to correct for a systematic error caused by an asymmetric covering of the grating.

15.5.3 Gap independence

In this section we derive the independence of ISPI phase from variations in the mask–substrate gap.

When the gap changes, the diffraction angles remain constant, but the path lengths and phase information in the subsequent diffracted beam vary, depending on the position of the grating from which the beam diffracts. The position on the grating, and hence the phase of a diffracted beam, will vary with gap. However, since the grating period is constant, the gap-dependent phase change of one beam is exactly balanced by an equal and opposite phase change of the corresponding diffracted beam, as depicted in Fig. 15.16. When these two beams interfere, their phase difference, and hence the fringe pattern, is identical, regardless of the gap.



15.16 Illustration of beams diffracted in the transverse plane at two different gaps (a) and (b). The phase of each beam diffracted at angle ϑ_1 will change, due to its sweep in position along the grating as the gap changes. However, the phase change of each beam will be equal and opposite, providing the grating periods are constant and the mask and substrate planes are parallel, resulting in an identical phase relation in the interfering beams.

As illustrated in Fig. 15.11, the phase of each beam diffracted by the substrate grating from either side is ϕ_1 and ϕ_2 at an initial gap. The phase of the beams diffracted at a slightly different gap are ϕ_3 and ϕ_4, where

$$\begin{aligned}\phi_3 &= \phi_1 + \Delta\phi \\ \phi_4 &= \phi_2 - \Delta\phi\end{aligned} \qquad [15.25]$$

resulting in

$$\phi_1 + \phi_2 = \phi_3 + \phi_4 \qquad [15.26]$$

and identical fringe patterns, regardless of the gap.

15.5.4 Low-numerical aperture (NA) long-working distance optics

Optical resolution, i.e., the ability to distinguish two closely spaced objects, can be characterized by the highest spatial frequency in a Fourier series composition or, alternatively, the minimum resolvable period in a grating, $p_{\min}$

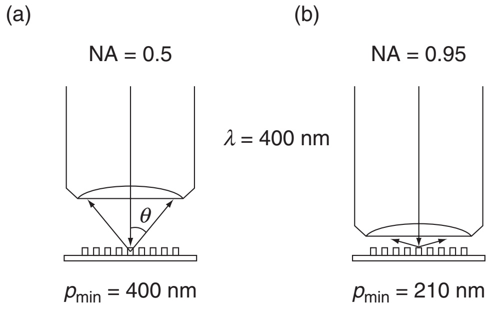
$$p_{\min} = \frac{\lambda}{2\mathrm{NA}} \qquad [15.27]$$

where λ is the wavelength of illumination and NA is the numerical aperture. NA is defined as

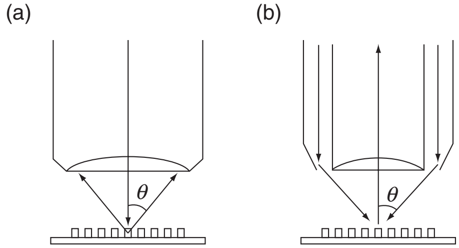
$$\mathrm{NA} = n\sin\theta \qquad [15.28]$$

where n is the index of refraction of the medium between the microscope objective and the imaged object, and θ is the maximum half-angle over which light can be collected by the objective lens (Fig. 15.17). Light can be collected by the lens in either a brightfield or darkfield configuration (Fig. 15.18). For visible light, in the range between 400 and 700 nm, a high-quality optical microscope has minimum resolution in air of $p_{\min} \approx 0.2$ µm. Resolution will be improved by increasing the index of refraction, n, in the optical path between the objective and the sample. For instance, oil immersion ($n = 1.4$) reduces $p_{\min}$ to ~0.14 µm.

In most optical metrology, high resolution demands high NA. However, since ISPI encodes position exclusively in low-spatial-frequency signals, which typically contain spatial frequencies one or two orders of magnitude lower than the physical gratings, position resolution is effectively decoupled



15.17 Lens geometries for two values of NA. The working distance of the lens decreases with increasing NA to capture the beams diffracted at larger angles. According to the Abbé theory of imaging, at least two orders must be captured to reproduce the fundamental grating period in the image plane of the microscope.



15.18 (a) Schematic of a conventional brightfield microscope. Illumination is at normal incidence. First-order diffracted beams are collected by the objective lens. (b) Schematic of a darkfield microscope. Illumination is axially symmetric and at oblique incidence. First-order diffracted beams return to the objective lens. Zero-orders reflect to the sides, and escape capture by the lens. The darkfield microscope is suitable for viewing samples with arbitrary feature geometries. ISPI uses a specific feature geometry that allows specialized illumination and viewing angles.

from microscope resolution. As a result, requirements on the numerical aperture of the microscope can be reduced considerably, from NA = 0.95 of a high NA microscope to NA = 0.06 of an ISPI microscope, without degrading position resolution.

A side effect of low-NA is that the working distance can be increased from a few millimeters for a high NA microscope to over 100 mm (WD = 110 mm in the current implementation) in the case of an ISPI microscope.

Another advantage of a low-NA microscope is a large depth of focus. Clearly, it is preferable to have a large depth of focus and obtain an alignment reading using a single image, as is required for continuous feedback control. In the current implementation, the ISPI microscopes have a 150 micron depth of focus, which is sufficient to observe alignment at nearly all practical mask–substrate gaps.

Low-NA optics has the further advantage of spatially filtering return signals. Since the ISPI microscopes subtend a small solid angle, with a very small aperture at a large distance from the mask and substrate, light randomly scattered from features on the substrate or mask will not be collected – with the exception of light scattered within an extremely narrow range of angles. This spatial filtering contributes to a low background light level and high SNR, but assumes that the microscope is at normal incidence to the sample. In the following section, we will take spatial filtering one step further, changing the viewing angle and excluding all reflected light from the ISPI image.

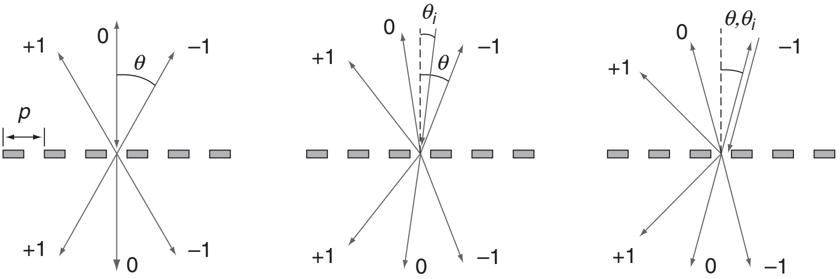
15.5.5 Littrow-angle viewing

In a conventional brightfield microscope, much of the light that returns to form the image comes from specular reflections, as indicated in Fig. 15.18a. A bright reflected component makes it difficult to obtain high contrast in moiré fringes. A step toward improved moiré fringe contrast is to use a conventional darkfield objective, in which illumination is at oblique angles, and with axial symmetry, as indicated in Fig. 15.18b. In such a darkfield objective, the zero-order beams are not collected, but instead traverse beyond the acceptance range of the objective lens. Consequently, the only light returning to the microscope results from diffraction and scattering from the substrate. However, such a conventional darkfield microscope is not ideally suitable for ISPI since it operates directly above the mask and would block any exposing radiation, such as UV flood exposure in imprint lithography.

We replace the conventional axial symmetry of the illumination and imaging paths with symmetry perpendicular to the incident plane, effected with a periodic structure, as illustrated in Fig. 15.19.

As the illumination angle increases away from the surface normal, backdiffracted orders increase in angle in the opposite direction. At a particular oblique angle, referred to as the Littrow angle, the incident and backdiffracted beams become collinear. The condition for the Littrow angle θ_L is described by

$$\theta_L = \sin^{-1}\left(\frac{\lambda}{2p}\right) \qquad [15.29]$$

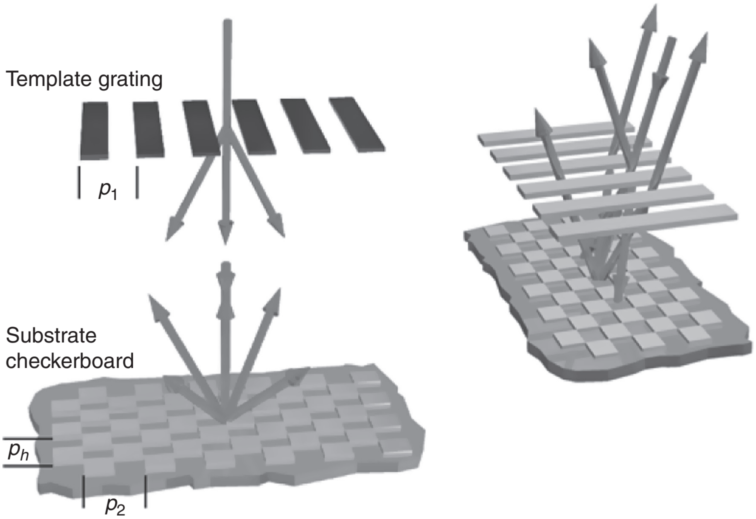


15.19 Diffraction from a grating of period p at various angles of incidence. In the first figure, the incident beam and zero-order reflection are collinear. As the angle of incidence increases away from the normal, the angle of one first-order diffraction decreases towards the normal. At a particular angle (called the Littrow angle) the first-order beam returns along the same path as the incidence beam. The result is that a single optic can be used for both illumination and imaging at an oblique angle, allowing a convenient and practical implementation of an ISPI microscope. Imaging at the Littrow angle is important to feedback-controlled alignment and gap control at any time, including during exposure. Littrow imaging also acts as a filter to prevent any light from returning to the microscope except that which is designed to backdiffract at a particular angle, thus eliminating most degrading influences of defects or particles.

where λ is the wavelength and p is the grating period. This period is also referred to as p_h, to distinguish it from p_1 and p_2. ISPI typically uses $p_h = 1$ µm and λ from ~650–690 nm, resulting in $\theta_L \sim 20°$.

We take advantage of Littrow-angle imaging to spatially filter the fringes that reach the image sensor. Light is emitted from a fiber optic positioned near the optical axis of the ISPI microscope and reflected downwards near the Littrow angle; light diffracted from the ISPI marks returns to the microscope at a similar angle, resulting in high efficiency and strong SNR. Figure 15.20 shows schematics of the diffracted orders from both mask and substrate marks in (a) normal incidence and (b) Littrow angle. It is apparent that the period p_h responsible for returning light at the Littrow angle is only on the substrate. Furthermore, p_2 and p_h can be adjusted independently, providing flexibility in the selection of the moiré fringe period and the viewing angle. Note that the zero-order beams traverse well outside of the acceptance angle of the microscope, resulting in high fringe contrast and high SNR. Since the returning beams contain only higher diffractive orders, an additional factor of two in fringe motion is obtained, compared with fringes viewed at normal incidence.

Another key advantage is that the Littrow-angle viewing allows alignment to be measured continuously, before as well as during lithography or any other operations on the objects.

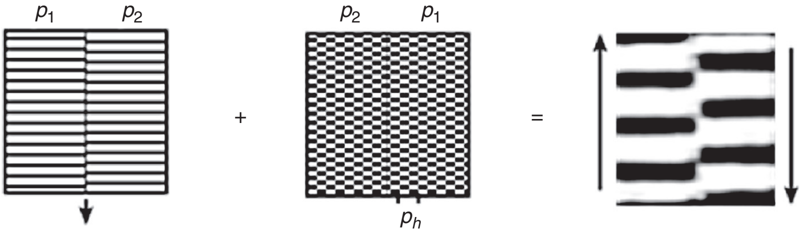


15.20 Schematic drawings of diffracted beams from grating and checkerboard marks on the mask and substrate. (a) Orientation of diffracted beams at normal incidence. Note that the checkerboard suppresses diffraction except along the diagonal directions. (b) Littrow-angle diffraction. Note that one pair of diffracted orders returns in the direction of the microscope, while the zero-orders and numerous other orders do not contribute to fringe formation. Note also that the period p_h directs the interference fringes towards the imaging optics at the Littrow angle, and can be adjusted independently of period p_2.

15.5.6 Integrated phase reference

A distinctive advantage of ISPI over other types of interferometers is the ability to analyze two displacement-sensitive signals at once, with the same sensor, and obtain a position measurement from the relative phase of the two measurements. The advantage of this method is that most error sources that affect one measurement will also affect the other measurement by the same amount. We refer to this property as 'mutual phase referencing.'

Encoding alignment in two matched fringe sets is a robust approach since the periodic structures have identical responses to etching and processing steps, by virtue of their nearly identical spatial frequencies. In addition, by encoding position in the relative phase of matched fringe sets, any perturbation caused by changes in the index of refraction in the beam paths will affect both fringe sets by equal amounts (assuming homogeneous media in the narrow ~400 μm beam width), thereby eliminating variations in the relative spatial-phase measurement. In short, both sets of fringes are affected by equal amounts by any reasonably homogeneous medium they traverse, resulting in immunity to variations in the index of refraction and other perturbations.



Mask: gratings

Wafer: checkerboard gratings

Interference fringes on CCD

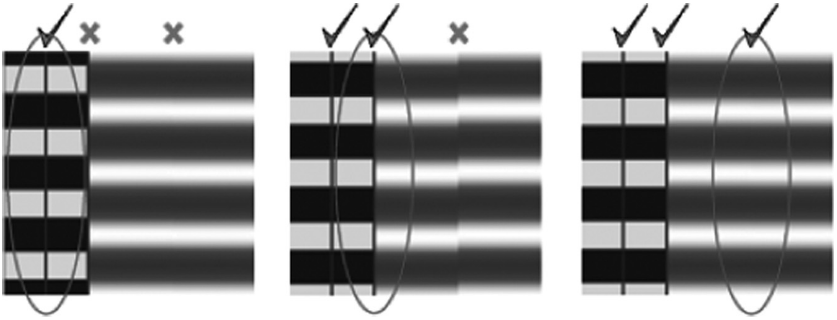
15.21 Alignment marks and interference fringes resulting from inverted grating periods. Motion of the mask grating results in magnified, counter-propagating motions of the interference fringes. Period p_h directs the interference fringes towards the imaging optics at the Littrow angle.

As indicated in Fig. 15.21, one fringe set is produced by superposition of p_1 over p_2, and the other fringe set is produced by superposition of p_2 over p_1. The two fringe sets move in opposite directions due to the inversion of the grating periods. Counter-propagating fringe motion doubles sensitivity, since an identical phase condition is repeated in half the displacement required with a fixed reference.

Any scheme that encodes information in the relative position of periodic patterns encounters positions at which the fringes appear indistinguishable from their starting position. Some means of resolving phase ambiguity, i.e., distinguishing between identical phases separated by 2π intervals, must be employed.

In ISPI, we take advantage of the ability to simultaneously analyze both *geometric* and *interferometric* features in the same image. Adjacent to the moiré fringe sets, we place small-area gratings, with their k-vectors parallel to both the incident plane and the moiré fringes. Each small-area grating backdiffracts directly to the microscope and is imaged as a bright spot. A set of such small-area gratings is referred to as a *bar array*. A bar array can be considered to be a 'fringe set' with zero displacement magnification. The period of the bar array is designed to match the spatial period of the moiré fringes. One bar array is on the mask, and another is on the substrate, in a non-overlapping position. Unambiguous alignment is achieved through a three-stage process, as illustrated in Fig. 15.22.

In the first stage, the relative spatial phase of the mask and substrate bar arrays indicates coarse alignment. In the second stage, the relative spatial phase between aligned bars and one moiré fringe set determines intermediate alignment. In the third stage, the phase between the moiré fringes indicates fine alignment. The same spatial-phase algorithm described in the previous section is used for each stage, with the appropriate phase-



15.22 Schematic of unambiguous phase alignment. On the left of the moiré fringes, a set of gratings (1 μm period in X) with periodic, vertical (Y) interruptions backdiffract periodic patterns of light to the ISPI microscope. The pattern has the same period as the moiré interference fringes (right). Unambiguous alignment consists of three steps: (a) Coarse alignment: Lining up the spatial phase of the bar patterns on the mask and substrate indicates coarse alignment from >30 μm to 30 nm. (b) Spatial phase between one set of bars and one set of fringes indicates intermediate alignment. (c) Fine alignment is determined by matching the spatial phase between two sets of interference fringes. The same phase algorithm is used in each case, with the exception of a scaling factor involved in the conversion of phase to displacement.

displacement conversion factor in each case. The displacement of the bars is given by $\Delta x_B = (\phi\ P_F)/(4\pi)$, where ϕ is the bar-bar phase and $P_F = 41$ μm. Minimum detectivity of the bar array is ~30 nm. Coarse position detection extends to at least one period of the bar array, and can be extended further by counting the number of bar array periods away from the center position. In this way, spatial-phase ambiguity is eliminated and the dynamic range extends from <0.05 nm to ~50 μm, or a factor of 10^6.

Although there is significant overlap between the detectivity of the bar–bar phase and the phase ambiguity of the fringes, the intermediate step involving bar-fringe phase provides effective assurance of the desired position: the bar–fringe phase is π-phase-shifted for every 2π-phase shift of the fringes, as illustrated in Fig. 15.22. Thus the most critical, 2π-phase ambiguity of the fringes is indicated by a π-phase shift between bars and fringes, providing a clear indication of the most likely type of alignment error.

15.6 Performance of ISPI

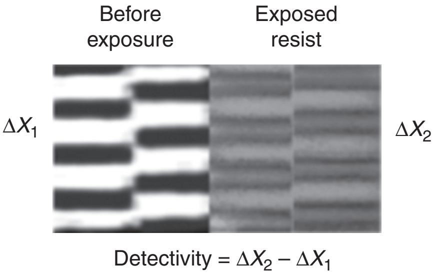
An experimental verification of the detectivity of ISPI compared the spatial phase detected by ISPI microscopes prior to exposure with the spatial phase of a moiré pattern in the resist after exposure and development. Before the exposure, a fringe pattern resulting from the interference of diffracted beams was observed by the ISPI microscopes. After the exposure, a moiré pattern between the patterned resist and the complementary pattern etched into the substrate was viewed at normal incidence by an ordinary optical microscope. Although the two types of fringes are observed with different microscopes with different viewing orientations, and are formed by different physical mechanisms, the spatial phase should be the same in both cases. The degree of agreement between the two moiré patterns is taken as an indication that the pattern transferred to the resist is highly correlated with the ISPI measurement and the ability of the fringes seen before exposure to successfully predict the critical figure of merit: the phase, and hence alignment, of patterns exposed in resist. These experiments were done with the mask in contact with the substrate to (a) avoid drift of alignment during exposure, and (b) to eliminate alignment artifacts due to the divergence of a point source.

The question of ISPI detectivity was separated from these issues simply by taking the difference of the spatial-phase measurements before exposure (in an ISPI microscope) and after exposure (in resist, viewed with a normal-incidence microscope). Of course, the absolute value of the measurements exhibited the distortions discussed above, but the difference of the spatial phases is the only quantity necessary to determine correlation between observed and actual alignment.

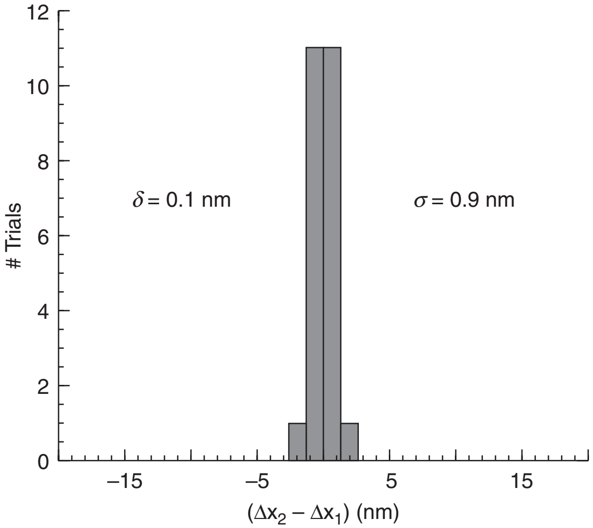
A final issue is how to measure with nanometer accuracy the placement of the patterns in the resist relative to patterns etched previously in the substrate.

With the above considerations in mind, we designed a special alignment test mask and conducted the X-ray exposure tests as follows. We wrote on a single X-ray mask that contained 96 of the $\{p_1, p_2\}$ alignment marks appropriate for ISPI, plus a variety of additional patterns. The alignment mark gratings had periods of 1.00 and 1.05 μm. A first X-ray exposure in contact transferred the mask patterns onto a silicon substrate, where they were subsequently etched in relief using reactive-ion etching. The substrate was then recoated with SAL 601 resist and placed back in the ISPI alignment apparatus. The X-ray mask was rotated 180° which, by virtue of its designed symmetry, brought $\{p_1, p_2\}$ gratings on the mask into superposition with $\{p_2, p_1\}$ marks on the substrate (i.e., a 180° rotation converted $\{p_1, p_2\}$ pairs to $\{p_2, p_1\}$ pairs). The mask was aligned and lowered onto the substrate, where it went into intimate contact. A lateral shift of ~100 nm occurred when the mask went into contact. The lateral shift is attributed to small inequalities in the height of the mesa surrounding the mask membrane. The shift occurred despite leveling the mask prior to bringing it into contact with the substrate, and varied by tens of nm between repeated contact cycles. After the mask was in intimate contact, we observed and recorded the phase disparity between the two halves of each mark, and converted that into a displacement, called ΔX_1. Following exposure, we developed the resist and measured the overlay of the $\{p_1, p_2\}$ resist patterns on the $\{p_2, p_1\}$ relief gratings in Si for all of the mark pairs, using a normal-incidence Leitz microscope. A representative example of both fringe pairs is shown in Fig. 15.23.

The measured displacement of the moiré image of resist on a silicon relief grating is called ΔX_2. We call the quantity $(\Delta X_1 - \Delta X_2)$ the 'disparity between detected and measured misalignment.' The alignment difference was found



15.23 (a) Interferometric fringe pattern observed through an ISPI microscope. (b) A moiré fringe pattern between exposed resist and a grating etched in a substrate, observed by a Leitz microscope.



15.24 ISPI alignment detectivity, showing a histogram of $(\Delta X_1 - \Delta X_2)$ for the 1 μm-period alignment marks. The difference of moiré fringes as detected in an ISPI microscope before exposure and fringes observed in resist with a different microscope at normal incidence after exposure and development. The X-ray exposure is done with the mask in contact with the substrate to eliminate the effect of point-source magnification. The mean difference was 0.1 nm and σ = 0.9 nm.

for 24 marks. The difference of these measurements is plotted in Fig. 15.24 for marks with 1.00 and 1.05 μm grating periods. Alignment as perceived by ISPI and alignment shown in the exposed pattern agreed, with a mean difference of $\delta = 0.1$ nm, with $\sigma = 0.9$ nm.

We believe this experiment proves that there is a correspondence between the misalignment, as detected by the ISPI optics and algorithm, and the misalignment as measured in exposed resist. We believe the results demonstrate that the ISPI scheme has a detectivity of misalignment better than 1 nm, with a standard deviation <1 nm.

It should be emphasized that the same marks are used for both measurements, yet the means of detection are quite distinct. There is a substantial difference between the two types of fringes, in both their underlying physical formation and the observation mechanisms. The ISPI fringes are viewed at an angle of 20° using the low-NA, low-magnification, darkfield ISPI microscopes with multiple laser line illumination, while the fringes in resist are viewed at normal incidence with a high-NA, high-magnification, brightfield microscope and white light. Despite these differences in the means of viewing the fringes, the disparities between the two forms of measurements were <1 nm.

15.7 Backside ISPI

Although frontside alignment marks are in nearly universal use, they are always subject to process-induced degradation, occupy valuable substrate area, and are necessarily of limited dimensions. In many circumstances, it is desirable to align an imprint template, a mask, or other structure to a substrate without dedicating substrate area to alignment marks, or requiring fabrication of new marks (and cumulative alignment errors) with every subsequent pattern level. In many applications, it is also important to maintain spatial coherence across large areas.

Previous attempts to use backside alignment marks employed visible wavelengths, and required customized wafer chucks with built-in optical paths to observe the position of the backside marks from underneath the substrate, while additional optical paths imaged mask and wafer-chuck alignment marks from the topside. Position detectivity was limited to the 1 μm level by conventional high-magnification visible-light optics, and practical overlay was constrained by distortions from the transfer optics in the wafer chuck.

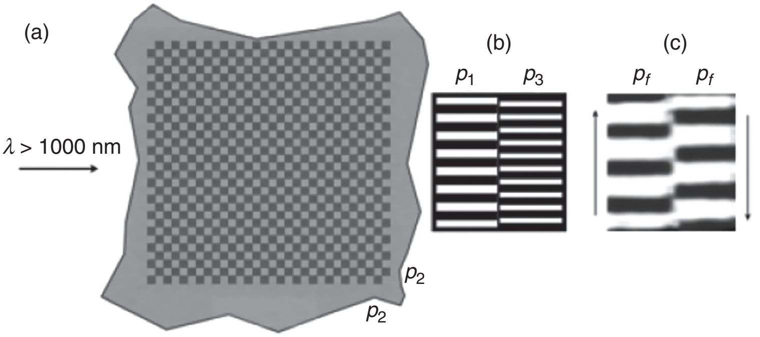
An alternative to imaging the backside using additional optics in the chuck is to employ infrared illumination, in the silicon transparency band. Unfortunately, use of infrared illumination with conventional optical microscopes will invariably reduce resolution, typically to the few μm range.

In contrast, since ISPI fringe formation depends upon the grating periods and not the wavelength of illumination, we can use infrared light with the same grating periods *without loss of resolution*. The restriction is the minimum grating period, which needs to be slightly larger with infrared light. To obtain diffraction, the minimum grating periods must be larger than the illumination wavelength, e.g., for illumination of 1065 nm, the grating periods could be ~1.6 microns. As long as the gratings can diffract at the illumination wavelength, the fringe period, moiré magnification all work as they do with visible illumination.

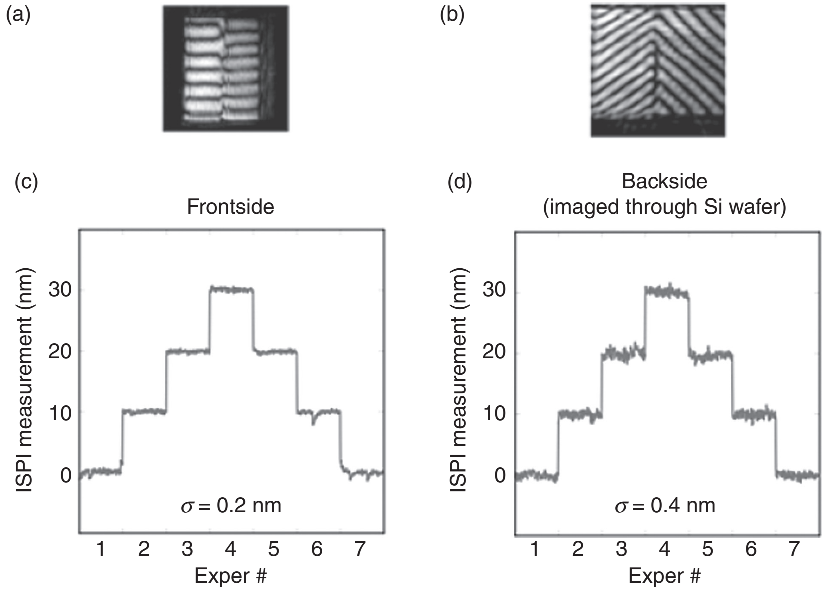
In this section, we therefore describe a variation of ISPI that embeds alignment marks on the backside of a double-side-polished substrate, providing a permanent, intrinsic coordinate reference, while requiring zero frontside footprint.

Infrared-based ISPI alignment is optimized by a variation on the mark design described in Section 15.2. The backside ISPI mark design (Fig. 15.25) consists of a p_2 period checkerboard on the substrate, with the same period in both orthogonal directions. Complementary to this backside checkerboard is a bifurcated grating mark on the mask surface, consisting of grating periods p_1 and p_3 that are respectively larger and smaller than the checkerboard period ($p_1 > p_2 > p_3$). Interference fringes translate in opposite directions in response to a displacement of template relative to substrate, in a manner identical to that observed with ISPI and visible illumination. In the case of substrate backside marks, the operating range is increased, and is limited only by the size of the substrate backside mark and the spatial coherence of the illumination. The backside mark can be increased to arbitrarily large dimensions to provide the basis of a substrate-wide coordinate system.

In an experiment, the detectivity of backside marks was compared with frontside marks. A double-side-polished, 100-mm diameter Si wafer (375 μm



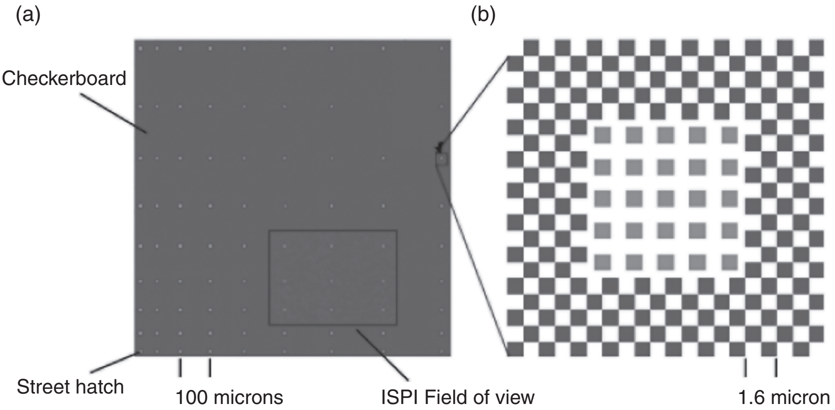
15.25 Backside ISPI marks variation including (a) the backside checkerboard with period p_2 in both orthogonal directions, (b) a $\{p_1, p_3\}$ grating set on the mask, and (c) resulting counter-propagating moiré fringe sets, observed when the marks are illuminated with substrate-transparent light (such as in the silicon transparency band >1 μm).



15.26 (a) CCD image of ISPI fringes using frontside marks. (b) CCD image of ISPI fringes using backside marks. (c) ISPI position measurements taken from frontside marks during 10 nm steps of a closed-loop piezo stage moving the substrate. (d) ISPI position measurements taken from backside marks during 10 nm steps of a closed-loop piezo stage moving the substrate. Note that position detectivity in the two cases is nearly identical, and has a standard deviation of 0.4 nm for the backside marks.

thickness) was used for the substrate, and a 38.1 mm diameter, 6.35 mm-thick fused silica optical flat was used for the template. A 300 nm layer of poly(methyl methacrylate) (PMMA) protected backside marks on the substrate. The illumination wavelengths were 660 and 1065 nm. Figure 15.26a and 15.26b show images of interference fringe sets generated by frontside and backside marks. Data obtained during piezo-controlled 10-nm displacements of the substrate are shown in Fig. 15.26c and 15.26d. It is apparent that the detectivity of the frontside and backside marks is essentially identical, and is maintained at a sub-nanometer level while viewing backside marks.

Spatial ambiguity over the large-area substrate mark is eliminated using small regions, or islands, consisting of a few periods of a grid that backdiffracts light directly to ISPI microscopes (Fig. 15.27). The islands are interspersed throughout the checkerboard at periods that increase monotonically in two dimensions. Back diffracted light from several islands is imaged in an ISPI microscope at the same time as the interference fringes, resulting in a pattern from which one can extract fine position (from the spatial phase of the fringes) and absolute position (from the spatial frequency of the islands). In



15.27 (a) Schematic of a region of a backside ISPI mark. The solid color region consists of a p_2 checkerboard, as shown in (b). Unambiguous position information is achieved using islands of backdiffracting, hatched grating regions, whose monotonically-increasing spatial frequency is compared with the moiré fringe frequency.

this manner, backside ISPI marks combine sub-nanometer detectivity with large-area coverage and suitability for multiple frontside alignment steps.

15.8 Conclusion and future trends

This chapter describes an alignment method referred to as ISPI. The central purpose of this method is to align objects such as a mask and substrate at the nanometer or sub-nanometer level, with measurements unaffected by many spurious influences, such as the refractive index of the medium in the beampath and process coatings covering the substrate alignment marks.

To meet these goals, ISPI encodes position in the spatial-phase disparity between a matched, adjacent pair of interferometric moiré patterns, viewed at oblique angles with low-NA microscopes. Phase referencing between the moiré patterns and design of the diffraction paths eliminates numerous environmental error sources. Oblique angle viewing avoids zero-order contribution to image formation, thereby providing high-contrast moiré fringes, high SNR, and doubling spatial frequency and sensitivity. The optics are removed from the exposure path, allowing alignment to be measured and corrected at any time, before or during lithography. Low-NA optics filters all spatial frequencies higher than the moiré frequency and permits a large working distance.

ISPI is independent of the illumination wavelength, to encourage simultaneous use of multiple wavelengths and avoid loss of alignment signals that can result from thin-film interference effects. Infrared illumination can also be employed to detect alignment information from marks on the backside of a substrate.

In the future, backside ISPI marks can be fabricated on entire wafers (up to 300 mm) with ~1 nm precision, using SBIL. Large-area fiducial grids can have multiple applications, including alignment and construction of multi-level photonic or electronic structures to the same zero-level alignment mark (using viewing openings in metal layers), wafer-to-wafer bonding with overlay at the 1-nm level, cross-tool referencing to an intrinsic coordinate grid that is unaffected by frontside processing, and wafer-wide position control of scanning probes. ISPI marks can be placed on both front and back substrate surfaces, and imaged simultaneously with one ISPI microscope, allowing *in situ* lateral and vertical measurement of thermal or stress-induced distortions. Further applications of backside ISPI could include sub-nanometer position referencing for electron beams or other charged particle beams.

15.9 References

1. SussMicroTec Inc., 228 Suss Drive, Waterbury Center, VT 05677 (www.suss.com).
2. B. Fay, J. Trotel and A. Frichet, 'Optical alignment system for submicron x-ray lithography', *J. Vac. Sci. Technol.* **16**, 1954 (1979).
3. G. Chen, F. Cerrina and Z. Wu, 'X-ray lithography two-state alignment system', *J. Vac. Sci. Technol. B* **7**, 1995 (1989).
4. G. Chen, J. Wallace, F. Cerrina, S. Palmer, B. Newell and J. Randall, 'Experimental evaluation of the two-state alignment system', *J. Vac. Sci. Technol. B* **9**, 3222 (1991).
5. G. Chen, J. Wallace, R Hachman, G. Wells, D Bodoh, P. Anderson, M. Reilly and F. Cerrina, 'CxrL aligner: An experimental x-ray lithography system for quarter-micron devices', *J. Vac. Sci. Technol. B* **10**, 3229 (1992).
6. H. Zhou and M. Feldman, 'Sub-nanometer alignment system for x-ray lithography', *J. Vac. Sci. Technol. B* **12**, 3261 (1994).
7. D.C. Flanders H.I. Smith and S. Austin, 'Theory of alignment monitoring by diffraction from superimposed dual gratings', *Appl. Phys. Lett.* **31**, 426 (1977).
8. S. Austin, H.I. Smith and D.C. Flanders, 'Alignment of x-ray lithography masks using a new interferometric technique – experimental results', *J. Vac. Sci. Technol.* **15**, 984 (1978).
9. M. Nelson, J.L. Kreuzer and G. Gallatin, 'Design and test of a through-the-mask alignment sensor for a vertical stage x-ray aligner', *J. Vac. Sci. Technol. B* **12**, 3251 (1994).
10. L. Raleigh, 'On the manufacture and theory of diffraction gratings', *Phil. Mag. Ser.* **4**, 47(310), 81–93 (February 1874) and 47(311), 193–205 (March 1874).
11. L. Raleigh, 'On copying diffraction-gratings, and on some phenomena connected therewith', *Phil. Mag. Ser.* **11**, 196–205 (1881).
12. J. Guild, *The Interference Systems of Crossed Diffraction Gratings – Theory of Moiré Fringes*, Oxford University Press, London, England (1956).
13. J. Guild, *Diffraction Gratings as Measuring Scales*, Oxford University Press, London, England (1960).

14. P. Konkola, Design and analysis of a scanning beam interference lithography system for patterning gratings with nanometer-level distortions, Ph.D. Thesis, MIT, Cambridge, MA (June 2003).
15. C. Chen, Beam alignment and image metrology for scanning beam interference lithography – fabricating gratings with nanometer phase accuracy, Ph.D. Thesis, MIT, Cambridge, MA (June 2003).
16. M.C. King and D.H. Berry, 'Photolithographic mask alignment using moire techniques,' *Appl. Optics*, **11**(11), 2455 (1972).
17. J.G. Goodberlet and B.L. Dunn, Deep-ultraviolet contact photolithography, *Micro Nano Eng.*, 1999, 11–12 (1999).
18. T.M. Lyszczarz D.C. Flanders, N.P. Economou and P.D. DeGraff, 'Experimental evaluation of interferometric alignment techniques for multiple mask registration', *J. Vac. Sci. Technol*. **19**(4), 1214 (1981).
19. A. Moel, E.E. Moon, R.D. Frankel and Henry I. Smith, 'Novel on-axis interferometric alignment method with sub-10 nm precision', *J. Vac. Sci. Technol. B* **11**, 2191 (1993).

16

Sidewall roughness in nanolithography: origins, metrology and device effects

V. CONSTANTOUDIS and E. GOGOLIDES,
NCSR Demokritos, Greece and G. P. PATSIS, Technological
Educational Institution of Athens, Greece

DOI: 10.1533/9780857098757.503

Abstract: One of the great challenges in next generation lithography is to print linear features with controllable sidewall roughness, which is usually called line edge/line width roughness (LER/LWR). The aim of this chapter is to provide an interdisciplinary approach to LER/LWR covering all related aspects. To this end, after a short introduction to LER/LWR concepts, it reports the basic findings of recent intensive research concerning the metrology and characterization, the material and process origins, and the device effects of LER/LWR. Both simulation and experimental results are presented, and emphasis is given to their comparison.

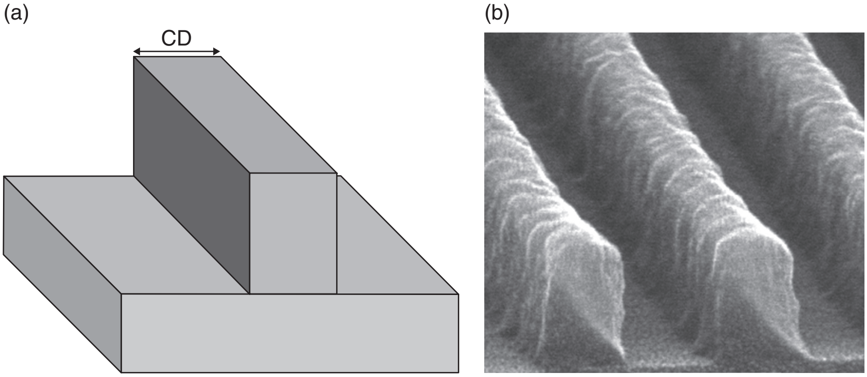
Key words: line edge roughness (LER), line width roughness (LWR), critical dimension uniformity (CDU).

16.1 Introduction

In this section, we introduce the term line edge roughness (LER) and explain its relationship with the sidewall roughness of lithographic patterns. Also, we discuss the motivations for studying LER and give arguments why this is more urgent at nanoscale patterns.

16.1.1 Sidewall roughness and line edge roughness (LER)

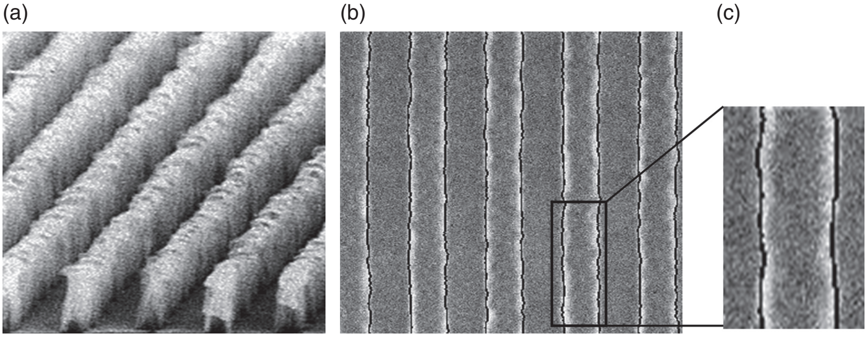
The aim of photolithography is to transfer a pattern from a mask to a light-sensitive polymer, the photoresist. The pattern can have different shapes, such as line-space structures, cylindrical or ellipsoidal holes, round corners, or more complex layouts. Undoubtedly, the simplest pattern that can be defined is a single linear feature (an isolated line) with length much greater than width. Usually, such lines are the first steps toward the



16.1 Ideal shape of a line structure on a substrate with smoothed sidewalls (a) and an SEM image of real resist line structure where the roughness of sidewalls is evident (b).

fabrication of transistor gates. This implies that the width of the line is a crucial dimensional parameter (called critical dimension (CD)) because it eventually defines the gate length in the fabricated transistor. The ideal shape in which the CD remains unaltered along the whole line is shown in Fig. 16.1a. However, following the famous Mandelbrot statement 'Clouds are not spheres, mountains are not cones, coastlines are not circles, bark is not smooth' (Mandelbrot, 1982), we observe that surfaces (natural and artificial) are also not absolutely smooth and flat but exhibit roughness (see Fig. 16.1b). In a linear feature similar to that shown in Fig. 16.1, roughness can refer either to the height sides of the line or to the sidewalls. The former modifies its thickness while the latter induces variations of CD along the line. Due to the importance of the CD, the sidewall roughness has attracted more research interest and is described in the literature as LER. The word 'edge' in LER has been chosen instead of the more descriptive 'sidewall', due to the appearance of a line in the top-down scanning electron microscope (SEM) images used widely in LER measurement. Owing to the edge effect in SEM measurements, the sidewalls are depicted in the top-down images as bright edges (see Fig. 16.2 and more details in Section 16.2). Many researchers prefer to use the term line width roughness (LWR), which refers to the variations of line widths (CDs) along the line from the nominal value, since this quantity describes more accurately the induced variations in the channel length of the transistor.

Quite surprisingly, despite its fundamental nature, the exact correspondence between the 3D sidewall roughness and the 2D LER is still an open issue. Up to now, it has been approached by simulating the formation of the top-down SEM images of resist lines through modeling of SEM e-beam interaction with resist, and comparing the input 3D sidewall roughness of the lines with the obtained 2D LER/LWR on the output SEM images



16.2 (a) 3D SEM image of a structure with lines and spaces, (b) the top-down SEM view of a similar structure with the detected line edges, and (c) magnification of a part of a line where the rough fluctuations of line edges (LER) is evident.

(Lawson and Henderson, 2010a; Li, 2008). However, a recent experimental finding may be critical to this issue: CD-AFM and CD-SEM measurements have revealed that the sidewalls of resist lines used in 193 nm and extreme ultraviolet (EUV) lithography exhibit strong anisotropy with stripes propagating from the bottom to the top of the line (curtain-like structures). Metrologically, this implies that the 2D obtained LER/LWR may be a sufficient and reliable representation of the real sidewall, since the latter has actually 2D structure (Constantoudis *et al.*, 2010; George *et al.*, 2010).

16.1.2 Motivations for studying LER

The main reasons for studying LER/LWR are first, the observation that it does not scale down with feature dimensions (CD) and thus the ratio LER/CD increases as the device sizes and lithographically fabricated features decrease (Rice *et al.*, 2003).

Second, it has been shown by atomistic simulations of a Metal-Oxide-Semiconductor Field-Effect Transistor (MOSFET) planar device that LER will become the dominant contributor to device degradation when CD < ~20 nm. It should be stressed that LER affects both single transistor performance and transistor variability (Roy, 2007; Reid, 2010a, 2010b). Similar studies have also revealed the critical importance of LER in the operation of the Fin Field Effect Transistor (FinFET), which has been recently introduced in the manufacturing of integrated circuits (Baraveli *et al.*, 2008; Yu *et al.*, 2009).

16.2 Metrology and characterization

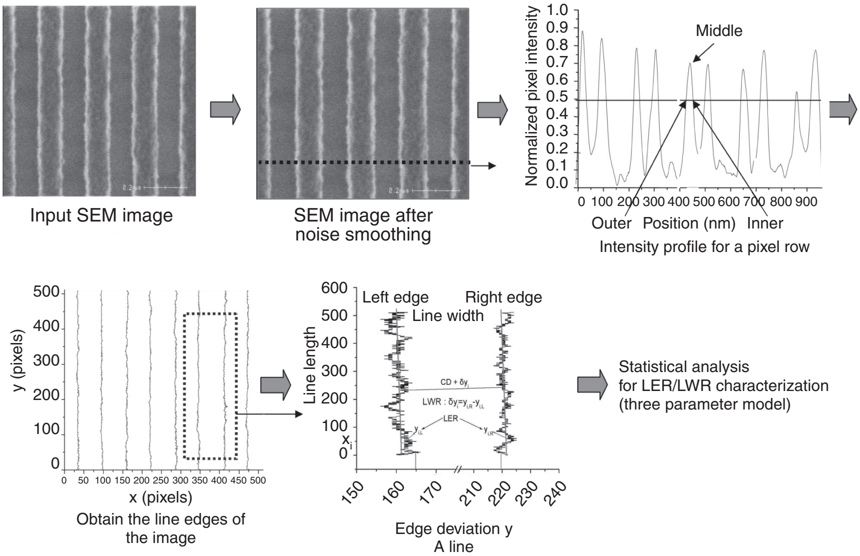
This section includes a short account of the most widely used measurement techniques and characterization methodologies in LER metrology.

16.2.1 Measurement techniques

The dimensional metrology of lithographic line-space structures has been usually performed with top-down CD-SEM images or scatterometry techniques. Both are widely used in CD measurements, with the first being more suitable for evaluating CD variations across lines, dies, or chips. LER/LWR is associated with local CD variations along a specific line, and therefore SEM has been the first choice for LER/LWR measurement and assessment. Scatterometry techniques and CD-AFM measurements have also been proposed and tested, but CD-SEM remains to date the workhorse metrology tool for sidewall roughness measurements in lithography. Below we will give a brief account of all metrological approaches, with more emphasis on CD-SEM and the subsequent characterization issues.

CD-SEM

Figure 16.3 shows the flowchart of the measurement and characterization of LER/LWR by means of top-down CD-SEM images. First we acquire a CD-SEM image of the line-space resist structure we intend to evaluate. Attention should be paid to the parameters of SEM measurement (magnification, resolution, frame number, dwell time) so that the charging effects and the sample damage caused by SEM electrons are under control. Recent


Input SEM image
SEM image after noise smoothing
Normalized pixel intensity
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Middle
0 100 200 300 400 500 600 700 800 900
Outer Position (nm) Inner
Intensity profile for a pixel row
y (pixels)
500
400
300
200
100
0
0 50 100 150 200 250 300 350 400 450 500
x (pixels)
Obtain the line edges of the image
600
500
400
300
200
100
0
Line length
Left edge
Right edge
Line width
CD + δy_i
LWR : δy_i=y_LR-y_LL
LER
y_LL
y_LR
x_i
150 160 170 210 220 230 240
Edge deviation y
A line
Statistical analysis for LER/LWR characterization (three parameter model)

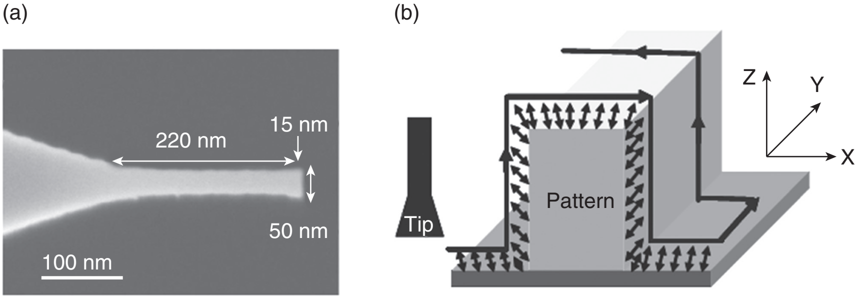
16.3 The steps followed for the extraction of line edges from an input top-down SEM image of line-space structure (Constantoudis *et al.*, 2003, 2004a).

works have shown that these effects may influence the obtained LER results (Patsis *et al.*, 2003; Yamaguchi and Yamamoto, 2008). Then, an image noise filter is applied to reduce the noise induced in the image by the measurement process. The choice of the filter parameters and type influence LER/LWR values, and thus they should be kept fixed when comparisons are attempted (Patsis *et al.*, 2003). Over the last years, the issue of noise effects on LER/LWR has been discussed widely, and several methods have been proposed for resolving it and providing noise-free LER/LWR measurements (Villarrubia and Bunday, 2005; Yamaguchi *et al.*, 2006; Hirawa and Nishida, 2009; Azarnouche *et al.*, 2011).

The smoothed image is subsequently analyzed to detect the edges of lines. This can be done either by putting a threshold in pixel intensities or by fitting the SEM signal perpendicular to the edge with polynomial curves (Bunday *et al.*, 2004). After the application of the edge detection algorithms, we get the coordinates of the pixels defining the detected edges of the lines included in the analyzed SEM image, and the statistical analysis of their roughness can be initiated. We will describe in more detail the parameters and concepts involved in Section 16.2.2.

CD-AFM

Initial attempts at applying conventional AFM to LER measurement included the cleaving of the resist line-space structure parallel to line direction and then the measurement of the sidewall roughness by the perpendicular to the sidewall oscillation of the AFM tip. Despite the useful results this method has provided, it is time-consuming, difficult to implement and limited to a specific portion of the line sidewall. To overcome these shortcomings, Martin and Kumar Wickramasinghe (1994) proposed and implemented a new AFM setup enabling the direct measurement of the surface topography of 3D structures. The key elements of this setup were the use of a tip with flared shape and the ability to perform oscillations in the plane (x, z) (see Fig. 16.4). These new developments, along with the advanced electronics for servo control, made this new AFM an appropriate tool for CD and LER/LWR measurements. First applications of the CD-AFM to LER metrology were limited by the large radius of the used tip (120 nm) (Orji *et al.*, 2005). More systematic measurements with smaller tip radius (50 nm) of the effects of acid diffusion length (ADL) and etching steps on LER/LWR were also performed and partially compared with the results of the more mature CD-SEM measurements (Foucher *et al.*, 2007; Pargon *et al.*, 2008). Similar trends vs CD were found, where the sidewall slopes were larger than 89°. However, systematically CD-AFM-measured LER was larger than that obtained from CD-SEM. This difference was attributed to the unclear relationship between the 1D CD-SEM representation of LER



16.4 Basic principle of the 3D AFM measurement: (a) the flared Si tip used for the measurements, and (b) the CD Mode scan principle (Foucher *et al.*, 2007).

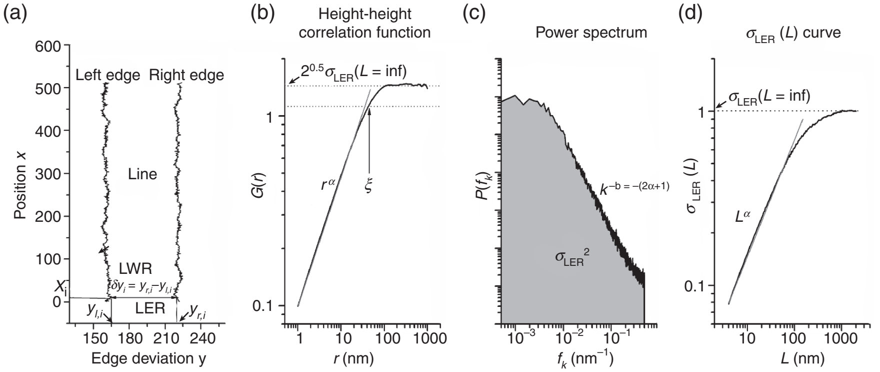
with respect to the more complete 2D sidewall topography extracted by CD-AFM, as well as to the tip size effects. Furthermore, the different contributions of noise effects in the two techniques may have an influence on the observed difference (Foucher *et al.*, 2007; Azarnouche *et al.*, 2011).

Scatterometry

Scatterometry techniques are widely used for CD measurements, due to their advantages (high throughput and non-destructivity), but have not yet been established for LWR metrology. Several relevant works have been published recently. The more systematic approach was undertaken by NIST, and was based on the use of small angle x-ray spectroscopy (SAXS) (Wang *et al.*, 2008). It was found that random LER is related to the ratio of the scattering intensities along the equatorial line of the diffraction spectrum when LER is present to the intensities of the ideal smooth (no LER) pattern. Comparisons with CD-SEM results showed superior evaluation of LER with scatterometry. This was attributed to the global character of scatterometry measurements and the concomitant inclusion of very low frequency fluctuations. Other approaches employ dark-field scatterometry, ellipsometry, angular resolved scatterometry, etc. (Brill *et al.*, 2010; Foldyna *et al.*, 2011; Shyu *et al.*, 2007). Modeling studies have also been performed, testing new techniques in simulating the effects of LER on scattered light such as the field-stitching method (Schuster *et al.*, 2009).

16.2.2 Characterization of LER/LWR

Let us assume that the line to be analyzed is known (has been measured) at N positions x_i, $i = 1, \ldots, N$ on both edges and $y_{l,i} = y_l(x_i)$ ($y_{r,i} = y_r(x_i)$) are the distances of the left (right) edge points from a reference axis on the left side



16.5 Detected edges of a typical resist line (a) along with the HHCF, (b) the power spectrum, (c) and the $\sigma(L)$ curve (d) of their roughness (Constantoudis *et al.*, 2004b).

of the line that is considered parallel to line direction (see Fig. 16.5a). LWR is the roughness of the differences (line widths) $\delta y_i = y_{r,i} - y_{l,i}$ while LER is the roughness of the $y_{l,i}$ or $y_{r,i}$. The following definitions of spatial roughness functions are written for the edge points y_i ($y_{l,i}$ or $y_{r,i}$) and refer to LER, but they also hold for the line widths δy_i and for LWR.

The most widely used metric for the quantification of LER/LWR has been the standard deviation of the edge points from their mean value (rms value) σ_{LER} :

$$\sigma_{\text{LER}}^2 = \frac{1}{N}\sum_{i=1}^{N}\left(y_i - \langle y_i \rangle\right)^2 \tag*{[16.1]}$$

where $<y_i>$ is the average of all y_i, $i = 1, \ldots, N$. Despite its widespread use, the rms value σ_{LER} suffers from some related drawbacks. First, it depends on the spacing between two nearby measurement positions d and the length L $(=(N - 1)d)$ of the edge included in the measurement process (He and Cerrina, 1998; Yamaguchi, 2003a; Constantoudis, 2003; Constantoudis, 2004a; Villarrubia, 2005). More critical is the dependence on L implying that actually σ_{LER} is not a single value but an increasing function of line length L. The 2007 edition of the International Technology Roadmap for Semiconductors (ITRS) recognized this dependence and dictated in the Lithography chapter that lines with lengths larger than 2 μm should be used in the estimation of σ_{LER}, based on the experimental observation that usually σ_{LER} saturates above this critical length (ITRS, 2007, p. 16). Second, the dependence of σ_{LER} is related to the spatial and/or frequency aspects of LER/LWR i.e. the spatial arrangement of the fluctuations of edge points or line widths along the line, which are overlooked in a single value of σ_{LER}, but may be critical for the effects of LER/LWR on device performance (Yamaguchi and Komuro, 2003b; Constantoudis, 2004b; Bunday, 2004). These spatial fluctuations can be quantified by three ways.

Correlation functions

The most commonly used metric of the correlations is the autocorrelation function, which for the case of edge points takes the form:

$$R(r = md) = \frac{1}{\sigma_{\mathrm{LER}}^2}\frac{1}{N-m}\sum_{i=1}^{N-m}\left(y_{i+m} - \langle y_i\rangle\right)\left(y_i - \langle y_i\rangle\right) \qquad [16.2]$$

where $d = |x_{i+1} - x_i|$ is the distance in the direction of the line between two neighboring measurements and $<y_i>$ the mean edge position. Due to the normalization prefactors in (Equation [16.2]) $R(r = 0)= 1$ while $R(r \to \infty) = 0$, since the correlations are expected to tend to zero at long distances. Examples of autocorrelation functions of resist LER/LWR can be found in (Yamaguchi and Komuro, 2003b; Bunday, 2004). Another correlation function used in LER characterization is the height–height (or height difference) correlation function (HHCF) $G(r)$, defined as (Constantoudis, 2003, 2004):

$$G(r = md) = \left[\frac{1}{N-m}\sum_{i=1}^{N-m}\left(y_{i+m} - y_i\right)^2\right]^{1/2} \qquad [16.3]$$

The HHCF is related by definition with $R(r)$ through

$$G^2(r) = 2\sigma_{\mathrm{LER}}{}^2[1 - R(r)] \qquad [16.4]$$

and is used to reveal the scaling properties of the analyzed data. Obviously $G(r = 0) = 0$ and $G(r \to \infty)=\sqrt{2}\sigma_{\mathrm{LER}}$. A typical $G(r)$ of resist LER is shown in Fig. 16.5b. At small r, $G(r)$ usually increases following a power law, which reveals the presence of fractality (or more accurately self-affinity) in LER. The exponent of the power law is usually called the roughness exponent α, and is related to the fractal dimension d of the edge through $d = 2 - \alpha$ ($0 < \alpha < 1$). Low (high) values of α correspond to edges or lines with enhanced (reduced) contribution of high frequency fluctuations to LER. As we can see in Fig. 16.5b, the power law behavior has an upper boundary, above which the HHCF saturates or oscillates randomly around the value $\sqrt{2}\sigma_{\mathrm{LER}}$. The upper boundary of fractality is defined by the correlation length ξ, which quantifies the average width of edge asperities. ξ may be defined by $R(\xi) = 1/e$ (or equivalently $G^2(\xi) = (1-1/e)2\sigma_{\mathrm{LER}}{}^2$), or by fitting $R(r)$ or $G(r)$ with an exponential or other suitable function to the neighborhood around $r = 0$. For $r > \xi$, the edge point correlations can be considered negligible. Therefore, the analysis of a typical HHCF of LER leads us to a three-parameter model for LER characterization consisting of the parameters σ_{LER}, ξ, α. The prerequisite for the validity and application of this model is the limited scale fractality of LER. In the case of the presence of some periodicity, the wavelength of the periodic repetitions should be added. In most real experimental lines, the correlation length is estimated in the range 20–50 μm, while the roughness exponent lies between 0.5 and 0.7.

Power spectrum

Fourier transform (FT) analysis is widely used in signal analysis for revealing the frequency content of a signal. Given that a line edge is actually a spatial signal, Fourier analysis can also be applied to the characterization of the spatial/frequency aspects of LER. The discrete FT of LER is defined as:

$$F(f_k) = \sum_{m=0}^{N-1} y_m{}^{e-2\pi ikm/N} \qquad [16.5]$$

and the more commonly cited power spectrum (PS) as the square of FT amplitude:

$$P(f_k) = \frac{2d}{N^2}\left|F(f_k)\right|^2 \qquad [16.6]$$

where $f_k = k/(Nd)$ is the frequency corresponding to the k-th component of the FT, and $d = |x_{i+1} - x_i|$.

A typical PS in log–log scale is shown in Fig. 16.5c. A plateau at low frequencies is followed by a power law decrease at high frequencies. Not surprisingly, the behavior is the inverse of that exhibited by HHCF (see Fig. 16.5b). The power law decrease is an indicator of the presence of fractal self-affine symmetry at high frequencies, corresponding to the short distance behavior of HHCF. The fractality is limited to high frequencies, whereas at low frequencies there is a uniform contribution from all frequencies resembling the white noise behavior of uncorrelated data. The exponent of the power law b is related to the roughness exponent α since $b = -(2\alpha + 1)$, whereas the correlation length ξ is related to the inverse knee frequency marking the passage from the plateau to the power law behavior. The amplitudes of PS are related to rms value σ_{LER} through the Parseval's theorem:

$$\sigma^2_{\mathrm{LER}} = \frac{1}{(N-1)Nd}\sum_{k=0}^{N/2} P(f_k) \qquad [16.7]$$

which means that $\sigma_{\mathrm{LER}}{}^2$ is equal to the area under the $P(f_k)$ curve.

In conclusion, the three-parameter model suffices to characterize the PS of LER obeying the fractal self-affine symmetry and exhibiting the form of Fig. 16.5.

The $\sigma_{\mathrm{LER}}(L)$ curve

In statistical theory, it is well known that the standard deviation of correlated data depends on the size of the data and this dependence is indicative of the nature and extension of correlations. Therefore, the estimation of the dependence of rms on the line length of the measured sample (the $\sigma_{\mathrm{LER}}(L)$ curve) may be considered the third method for the description of the spatial characteristics of LER (He and Cerrina, 1998). For LER exhibiting fractal self-affine behavior, the $\sigma_{\mathrm{LER}}(L)$ curve has a form similar to that of HHCF, but with different scales (see Fig. 16.5d). A power law growth degrades gradually as line length L increases, and at large L (>2 μm) saturates to its final (length-independent) value. It can be shown that the exponent of the power law at low L is the same as that of the HHCF (i.e. the roughness exponent), while the knee value at which saturation starts is called the rms-correlation length L_C. Numerical calculations with model edges and experimental results have shown that $L_C \sim 10\xi$ (Constantoudis *et al.*, 2004b). The meaning of the $\sigma_{\mathrm{LER}}(L)$ curve is that the intrinsic-to-an-edge rms roughness decreases when the length of the edge is reduced below the rms-correlation length (~100–500 nm), and that this reduction is controlled by the spatial LER parameters ξ and α. Consequently, the form of the $\sigma_{\mathrm{LER}}(L)$ curve can be fully characterized by the three-parameter model obtained from the study of the HHCF and PS. We will come back to this curve and its importance when we discuss the LER effects on transistor performance.

LER and LWR

The above spatial and scaling analysis of LER, along with the related definitions and formulas, can be transferred to LWR replacing y_i by δy_ι. For uncorrelated left and right edges exhibiting similar scaling behavior, it can be shown that $\sigma_{\mathrm{LWR}} = \sqrt{2}\sigma_{\mathrm{LER}}$ while the correlation length and roughness exponent remain almost unaltered. More information about the relationship of LER and LWR can be found in Constantoudis *et al.* (2003); Bunday *et al.* (2004).

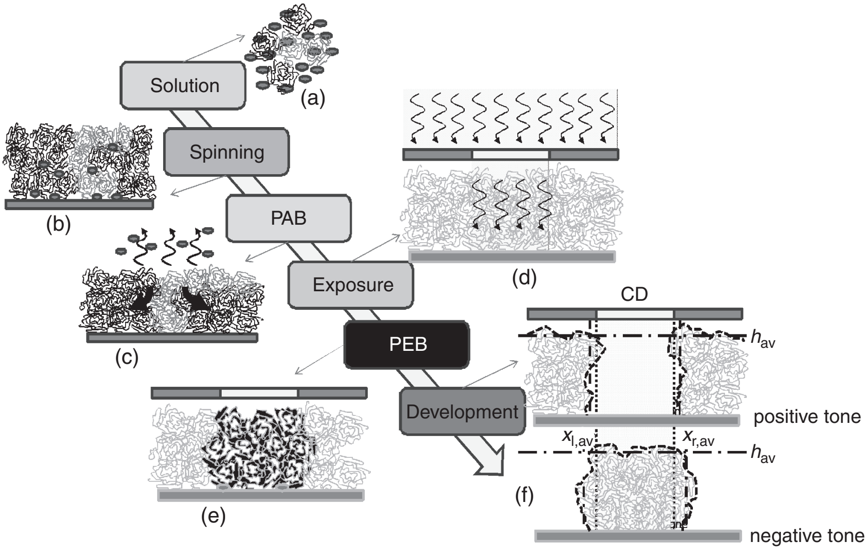
Experimental measurements and fractal self-affinity of LER/LWR

The three-parameter model introduced in the previous sub-sections is based on the assumption that LER/LWR exhibit fractal self-affinity symmetry for a limited range of scales, i.e. the edges and lines remain invariant under anisotropic scaling transformations for this range of scales. This assumption was verified by thorough experimental measurements (Leunissen *et al.*, 2005), which showed that indeed experimental LER/LWR from many resists exhibits scale-limited fractal self-affinity demonstrated by the power law parts in HHCF, PS and $\sigma(L)$ curve. Scale-limited power laws, especially in PS, have also been observed in the LER/LWR analysis from other researchers, and therefore can be considered a generic property of the conventional 193 and 248 nm resists as well as EUV resists (Yamaguchi *et al.*, 2003a; Bunday *et al.*, 2004; Lawrence, 2004; Larusso *et al.*, 2006; Hirawa and Nishida, 2010).

In summary, the majority of LER/LWR measurements come from top-down SEM images after application of in-line or off-line analysis to extract the edge points and/or line widths. Following edge detection, they usually calculate the rms value of the edge points or line widths, and if they aim at a more thorough study they proceed to the estimation of the PS. More rarely is the HHCF or $\sigma(L)$ curve calculated. Through these functions, the correlation lengths and roughness exponent are extracted by applying appropriate thresholds and estimating the exponents of the power laws respectively. Also, model HHCF and PS including the LER/LWR parameters have been used to fit experimental data and obtain the values of the parameters, while recently a bootstrap approach along with a modified variogram model have also been proposed (Patel *et al.*, 2010).

16.3 Process and material effects: modeling and simulation

LER is generally affected by lithographic processing steps and used materials. Recently, and in the context of EUV lithography, some discussion on the system-level effects has begun, with the main emphasis on mask roughness. Lithography incorporates several process steps as prior to etching shown qualitatively in Fig. 16.6, all of which act on the photoresist (see also in this book Smith, 2012). Material effects on LER are not easily quantified, because of photoresist complex structure and composition, and also because material effects are convoluted with process and system-level effects. In the case of chemically amplified resists, the resist is typically composed of a polymer matrix, a photoacid generator (PAG), and a quencher.



16.6 Qualitative depiction of the various lithographic process steps.

Both modeling and experimental work have been undertaken to better understand and control the impact of mask roughness, process conditions, and material properties on LER. In the present section, we will review briefly the modeling (analytical and numerical) approaches, while in the next one (Section 16.4) we will summarize the important experimental findings with reference to modeling predictions.

16.3.1 Analytical modeling

The first attempt at deriving an analytic formula incorporating the main dependencies of LER was made by Gallatin (2005). The derivation was based on a model of LER formation which included (a) random photon absorption and acid generation (photon shot noise), (b) acid diffusion during the post exposure bake (PEB) step, and (c) development of the deprotected fraction of resist using criteria inspired by the critical ionization model (Tsiartas *et al*., 1997). The model ignores inhomogeneities in material concentrations (resist, PAG, quencher) and dissolution rates, as well as the details of the reaction–diffusion kinetics. Also, it does not consider the molecular structure of the resist, and therefore molecular weight or polymer structure effects cannot be taken into account. The main outcome of the modeling is the analytical understanding of the trade-off between resolution, LER, and sensitivity (RLS trade-off), which guided recent resist research, especially in EUV lithography. Further extension of the analytical approach has been made by Mack (2010), taking into account the inhomogeneities in chemical concentration and the reaction–diffusion kinetics utilizing the concept of the von Smoluchowski trap for the polymer deprotection reaction.

16.3.2 Numerical simulations

Although analytical modeling approaches illuminate the fundamental dependencies of LER and provide straightforward interpretations of some of the experimental trends, the microscopic details of the resist and the relevant physicochemical processes can be captured only in numerical models simulating LER formation. This is especially true for the acid generation and kinetics during PEB, and the development process in which resist structure may play a significant role. Indeed, most proposed numerical models focus on the detailed simulation either of the acid reaction–diffusion kinetics or of the development of deprotected resist. All of these can be characterized as mesoscale, since microscopic models e.g. molecular dynamics are still limited to simulation scales of only a few nanometers, with long computational times. Table 16.1 lists the majority of the published simulation approaches for LER formation, and gives information about the modeling of the main lithographic steps and the predictions for LER (the models are referred to by the first authors of the related papers). For the sake of completeness, we also include in Table 16.1 the analytical approaches reported above. The predictions of the models (if any) are also referred to in the experimental results concerning material and process effects on LER reported in the following section.

16.4 Process and material effects: experimental results

In this section, we continue the discussion about the process and material effects on LER and summarize briefly the experimental results. Emphasis is given to the comparison with the modelling predictions

16.4.1 Mask roughness effects

In 193 nm lithography, the mask roughness is not considered a major contributor to resist LER due to the filtering out of the high and medium frequency LER fluctuations by the optical projection system. However, the effects of mask LER on aerial image contrast may be an issue, and have been studied in some papers (Mack, 2009; Garidis *et al.*, 2012).

In EUV lithography, in addition to mask LER, the impact of mask multilayer surface roughness should also be considered, since it leads to random

Table 16.1 Analytical and numerical models of LER formation: assumptions for the main lithography steps, predictions and some comments (the acronyms with asterisk are explained at the end of the table)

	Exposure	Material	Acid generation and kinetics	Development	Predictions	Comments
Gallatin (analytical) (Gallatin, 2005)	PSN[a] and ILS[a] included	Randomly distributed PAG[a] in a homogeneous continuous resist medium	Modeling of the acid diffusion around its generation point	Simplified version of CIM[a]	LER dependencies on dose, ADL, and ILS (RLS[a] trade-off), PS of LER	Provided quantification of RLS trade-off, successful comparison with experimental results for EUV
Mack (analytical) (Mack, 2010)	PSN[a]	Inhomogeneous chemical concentrations	Detailed diffusion-reaction kinetics with Smoluchowski's model		LER vs ADL	
Lawson (Lawson and Henderson, 2010b)	Cosine aerial image intensity variation	3D cell representation of resist. No cell-to-cell chain connectivity is considered. PAG[a] and quencher loading and aggregation included.	Dill C equation for the probability of acid generation, Kinetic Monte Carlo modeling of acid kinetics on resist lattice	Kinetic Monte Carlo approach based on a cell dissolution rate determined by the enclosed protecting groups	LER vs Dose-to-size, PAG and quencher base loading, ILS[a], PAG[a] aggregate size. Good correlation of LER with polymer protection gradient for homogeneous PAG distributions	Applied mainly to molecular resists, small simulation volume
Saeki–Kozawa (Saeki *et al.*, 2008)	Gaussian e-beam (radius = 1 nm) exposure: 3D electron trajectories in resist simulated with commercial software.	Continuous representation with acid generators and quenchers as spheres	Proton and counter anion distribution calculated through an ionization model for CAR[a] and Monte Carlo techniques	Mack's dissolution model (Mack, 1987)	LER *vs* exposure dose, development time, linewidth, quencher concentration. Full spectral analysis for the detection of low frequency LER causes.	Applied to e-beam and EUV lithography, Small simulation volume
Morita (Morita and Doi, 2010)	Ideal aerial image	Coarse-gained representation of polymers with dissipative particle dynamics method	A particle is converted to unblocked by the reaction contact with PAG	Removal of unblocked particles	LER vs polymer chain dynamics, development and rinse processes, number of block copolymers	

Biafore (Biafore *et al.*, 2010)	PSN[a] included	Lattice with atomic nuclei and a sea of electrons, random PAG[a] and quencher	Photoelectron scattering and Smoluchowski's model for reaction-diffusion of generated acids with PAGs and quenchers taken into account	No reference to development modeling	LER *vs* dose, focus and pitch. Also PAG[a] concentration effects on Esize and LER trade-off	
Schnattinger (Schnattinger and Erdmann, 2008)	Use of calibrated continuous models	3D cell representation of polymer, self-avoiding random polymers with chain structure and PAG[a] loading included	Dill C equation for the probability of acid generation, Kinetic Monte Carlo modeling of acid kinetics on resist lattice	CIM[a] with experimental calibration	LER (rms) *vs* development time, average polymer chain length, type of polymer chains, AIC[a], ADL, quencher concentration, exposure dose	
Patsis [Patsis *et al.*, 2005, Patsis *et al.*, 2010)	Ideal aerial image or imported from commercial software	3D cell representation of polymer, self-avoiding random polymers with chain structure and PAG[a] loading included	Acid generation through a threshold of a uniform probability. This threshold corresponds to the exposure dose. Monte Carlo modeling of acid kinetics on resist lattice	CIM[a] in real time, or fast projection in time	LER *vs* average polymer chain length, type of polymer chains, PAG concentration,. Also included effects of acid diffusion range on the polymer chain length impact on LER/LWR, full spectral analysis and inclusion of secondary electron diffusion in EUV	Used to study effects of molecular weight, chemical inhomogeneities. Also involved in scaling analysis studies.
Philippou (Philippou *et al.*, 2007)	Fixed AIC[a] from commercial s/w (c~0.8 with 9% flare)	3D cell representation of polymer structure obtained by MD[a] and self-avoiding random polymers	Stochastic modeling utilizing the cell resist description and specific reaction rates	Cell implementation of CIM[a] (no chain structure taken into account)	LER *vs* polymer length, PAG[a] concentration, CD and diffusion coefficient. Also correlation length and roughness exponent *vs*. polymer size	Small simulation volume, overestimation of high frequency LER.

[a] PSN = photon shot noise. ILS = image-log-slope, CIM = critical ionization model, RLS = resolution, LER = sensitivity, PAG = photoacid generator, AIC = aerial image contrast, CAR = chemically amplified resist.

phase variability in the reflected beam (speckle) and, hence, image plane resist LER (Naulleau *et al.*, 2010). The experimental characterization of these effects can be done by the 'exposure to exposure correlation method' proposed by Naulleau (2009). PS analysis of experimental results with correlated and anticorrelated line edge modulations of EUV resists revealed that both low and high frequency roughness were transferred to resist LWR on wafer (Pret and Gronheid, 2011). Modeling studies have shown that the spatial characteristics of the mask contribution to LER cannot be distinguished from the fractal behavior of resist-induced LER when stochasticity is included (Gallatin and Naulleau, 2011).

16.4.2 Process effects

Aerial image contrast

Aerial image is the light intensity distribution from the exposure tool incident upon the wafer. The 'steepness' of the aerial image at the edge of the patterned feature plays a large role in resist pattern quality. The closer an aerial image is to a step function in intensity, the better the final resist image is. The aerial image parameter that quantifies profile 'steepness' is the image-log-slope (ILS) defined by:

$$\mathrm{ILS} = \frac{1}{I(x_0)} \left. \frac{\partial I(x)}{\partial x} \right|_{x_0} \qquad [16.8]$$

where $I(x_0)$ is the image intensity at the line edge position x_0.

LER and LWR (i.e. their rms values) are generally observed to be inversely proportional to image log-slope (Shin *et al.*, 2001; Pawloski *et al.*, 2004):

$$\mathrm{LER, LWR} \propto \frac{1}{\mathrm{ILS}} \qquad [16.9]$$

This behavior has been predicted by both analytical modeling and numerical simulation (Gallatin, 2005; Schnattinger and Erdmann, 2008; Lawson and Henderson, 2010b).

Focus

The aerial image quality steadily decreases with both positive and negative defocus. Both CD-SEM and CD-AFM measurements indicate that LER and LWR values increase as the image moves out of focus (Foucher *et al.*, 2007; Kudo *et al.*, 2007).

Exposure dose

In optical lithography, exposure dose is related to the number of photons impinging on the resist film per unit surface, and involves the notion of stochastic variation or shot noise. Shot noise obeys Poisson statistics, and therefore the dose fluctuation is proportional to $1/\sqrt{N}$, where N is the number of photons incident upon a given resist area. Obviously, the importance of shot noise increases when N is small (highly sensitive resists, EUV lithography). Several experiments have shown that LER follows the exposure fluctuations and goes down with $1/\sqrt{E}$ where E ($\propto N$) is the radiation intensity (Brainard *et al.*, 2004).

The shot noise effect on LER/LWR has been captured from almost all models discussed above (see for example Patsis *et al.*, 2001; Gallatin, 2005; Schnattinger and Erdmann, 2008; Lawson and Henderson 2010b).

Post exposure bake (PEB) and acid diffusion length (ADL)

PEB is the main process factor for varying ADL, by varying either its temperature or duration. Other factors include the size of the catalytic acid molecule, the glass transition temperature (T_g) of the matrix resin, and the residual solvent content to achieve control of ADL. It is generally accepted and experimentally verified that for sufficiently large pitch the increase of ADL leads to LER/LWR reduction, due to the smoothing of spatial dispersion of photochemical events caused by shot noise.

However, it has been shown that for very large diffusion lengths relative to the pitch, acid diffusion causes degraded chemical contrast and exposure latitude, and consequently increased LER/LWR (Steenwinckel *et al.*, 2006).

The experimental trend of ADL in LER has been captured by the analytical approach developed by Mack (Mack, 2010), which also predicted that the optimum ADL increases with the deblocking reaction capture range (radius of von Smoluchowski trap). The same trend has also been predicted by some of the numerical simulation approaches (Schnattinger and Erdmann, 2008).

Resolution, LER and selectivity triangle

The abovementioned dependencies of LER on ILS, exposure dose, and ADL (for relatively small ratio ADL/pitch) can be expressed mathematically as:

$$\mathrm{LER} \sim \frac{1}{\mathrm{ILS} * \mathrm{ADL} * \sqrt{E}}$$

If we put together on the left-hand side the resist-dependent factors we get

$$\mathrm{LER} * \mathrm{ADL} * \sqrt{E} \sim \frac{1}{\mathrm{ILS}}$$

i.e. for fixed illumination conditions (ILS), we have a trilateral trade-off between LER, sensitivity or photospeed (*E*), and resolution (ADL) (Steenwinckel *et al.*, 2005). The reduction of two of these parameters comes at the cost of increasing the third. Since EUV resists should have low LER, high sensitivity (small *E*), and good resolution (small ADL), the so-called RLS triangle has been a big concern, and several papers have been devoted to proposing strategies for understanding and overcoming its limitations, such as the addition of acid amplifiers, synthesis of molecular resists, and the use of polymer-bound PAG polymers.

A theoretical framework for understanding the RLS trade-off has been provided by the analytical modeling of Gallatin (Gallatin, 2005). He obtained an analytical formula for the scaling dependencies of LER on shot noise (exposure dose *E*), ADL, and ILS:

$$\underset{\underset{\text{LER}}{\uparrow}}{\sigma_{\text{LER}}} \approx \left(\frac{I}{\partial I}\right)_{\text{edge}} \sqrt{\frac{T}{\bar{\rho}_{\text{PAG}}\,\alpha Q v \underset{\underset{\text{Dose}}{\uparrow}}{E}\,{}_{\text{size}}\underset{\underset{\text{Blur}}{\uparrow}}{R}{}^{3}}}$$

which was able to predict the basic scaling dependencies of LER on ILS, dose, and ADL (PEB temperature or time). Furthermore, the model can also reproduce the spatial characteristics of LER quantified in power spectral density (PSD) curves with appropriate determination of the ADL (Gallatin *et al.*, 2008).

Development conditions

Development time does not seem to have significant influence on LER/LWR (Kudo *et al.*, 2007), as also predicted in (Schnattinger and Erdmann, 2008). On the other hand, development concentration has been shown to affect LER, but the results are controversial, implying that these effects are strongly dependent on the resist chemistry (Kudo *et al.*, 2007).

16.4.3 Material effects

Polymer resist effects

One of the initial hypotheses for the explanation of LER was the formation of polymer aggregates (large groupings of polymer molecules) on the sidewalls of developed resist lines (Namatsu *et al.*, 1998). These aggregates with size ~20–30 nm have been shown to form during development because the polymers surrounding them dissolve faster than the aggregates do, due to differences in polymer density. Despite the first attempts at relating LER to the size of these aggregates and, hence, molecular weights, experimental measurements showed that the size of the polymer chain may be relevant to surface roughness, but its relationship with LER is more complicated since in the formation of the latter the ILS profile is further involved (Cutler *et al.*, 2003). The line of thought that polymers with smaller molecular weight will exhibit lower LER led to the proposal of molecular resists for new generation lithographies with hopefully improved RLS trade-off. Although these resists have not yet shown up till now any impressive reduction of LER/LWR, several proposals for both positive and negative tone molecular resists have been published, exploiting the advantages such as the fact that their synthesis and purification can be precisely controlled to create a mono-disperse resist with well-defined structure and properties.

Photoacid generator (*PAG*)

Preliminary studies have shown that increase in PAG concentration tends to reduce LER/LWR (Pawloski and Nealey, 2002; Gronheid *et al.*, 2008). However, it has been observed that the most commonly used ionic types of PAG tend to aggregate and phase separate (segregate) from the resist to some degree, leading to local inhomogeneities and thus increased LER. To overcome this problem, the novel concept of anchoring the PAG to the lithographic resist has been proposed and implemented with success. Several works have shown that these polymer-bound PAG resists exhibit better performance than the conventional resists (with blended PAG molecules) with respect to RLS trade-off limitations (Lee *et al.*, 2007). The principal role of PAG segregation has also been revealed by Fedynyshyn *et al.* (2009), who studied experimentally the surface roughness evolution of various open-surface resist films during development in order to isolate the material structure effects on resist roughness (the so-called innate material roughness).

Quencher

Chemically amplified resists contain also a small amount of basic species called quenchers, used to neutralize acid sites. It has been shown that moderately increased base loading causes a significant LWR reduction which has been attributed to the increased chemical gradient in blocked polymer after the PEB process (Michaelson *et al.*, 2005; Anderson *et al.*, 2008). On the other hand, very high quencher concentrations result in resist line degradation and increase in LER/LWR (Michaelson *et al.*, 2005).

Resist–substrate interactions

It has been reported that substrates may have a significant impact on LER, and the possible causes for these LER effects include surface morphology and substrate interactions, as well as residual reflectivity (Kudo *et al.*, 2007). As resist thickness goes down in NGL, resist–substrate effects are expected to play a more important role in LER formation. For example, it was recently observed that resist sidewalls after development exhibit anisotropy in the form of perpendicular-to-line-direction striations (Constantoudis *et al.*, 2010; George *et al.*, 2010). Furthermore, systematic investigations of organic and inorganic underlayers in EUV resists revealed the critical role of the coefficient of thermal expansion (CTE): low CTE underlayers exhibit better adhesion and consequently LER in thin film EUV resists (Higgins *et al.*, 2011). It is noteworthy that none of the proposed analytic or numerical models for LER formation can capture the formation of post-lithographic striations probably since they ignore the resist–substrate effects.

16.4.4 Post-processing treatments for reducing LER

Besides the conventional strategies for reducing LER/LWR (optimization of aerial image during exposure, ADL control during PEB, and development optimization), post development treatments have also been applied. They have the advantage of separating LER reduction from the other requirements that resists should meet. These post-processing treatments may be either chemical (rinse, UV vapor smoothing, ozonation, solvent treatment) or physical (ion implantation, e-beam curing, laser annealing) or physicochemical (plasma etching/trimming), and can cause LER reduction up to 40% (Chandhok *et al.*, 2008; Pret and Gronheid, 2010). However, the side effects that these treatments may have should be further considered (CD, sensitivity, etch resistance), and the survival of LER reduction after pattern transfer should be verified. First investigations have shown that smoothing after litho does not guarantee reduced LER after etching, and therefore more systematic studies are needed (Foubert *et al.*, 2011).

16.4.5 LER during pattern transfer

It is true that what affects device performance is not the resist LER but the sidewall roughness of the transferred feature on the Si substrate. The latter (substrate LER) depends on the resist LER and chemistry, the stack structure beneath the resist, and the etch process conditions. First experimental studies have shown that LER of the etched feature may be smaller than the resist LER, and also that substrate sidewalls exhibit anisotropic curtain-like morphology with vertical striations, which are induced by the anisotropy of the etching process (Goldfarb *et al.*, 2004). Reduction of LER has been confirmed with new experiments by Pargon *et al.* (2008), who used a new CD-AFM microscope for a direct measurement of the whole sidewall roughness, and Pawloski *et al.* (2006) for a range of initial resist LER values. The model studies can be separated into two categories. In the first, simple etching models are employed which take into account the ion shadowing by detailed models of resist LER, and predict the anisotropy of etched sidewall and the reduction of LER (Kokkoris *et al.*, 2009; Constantoudis *et al.*, 2010). In the second, the focus has been on the plasma–resist interaction, and they demonstrate the formation of perpendicular and parallel ripple formation on planar surfaces as a function of ion bombardment incidence angle as well as the smoothing and/or roughening of resist masks during etching (Guo and Sawin, 2009). Also, the plasma-induced surface modifications of blank resist films have been studied experimentally by various groups (Oehrlein *et al.*, 2010). They proposed that the roughness in resist surfaces is induced by the stress effects caused by the different elastic modules between a top graphitized layer (created by plasma ions) and a layer (modified by plasma vacuum ultraviolet (VUV) photons) beneath it. Finally, the key role of lactone groups in 193 nm resists and their removal during etching has been discussed (Pargon *et al.*, 2009; Soda *et al.*, 2009).

16.5 Impact on device performance

The impact of LER/LWR on device performance has been studied mainly for conventional planar MOSFETs and Fin Field Effect Transistor (FinFET) architectures. Herein, we will be limited to the planar MOSFET, while discussions and results about the LER/LWR effects on FinFET performance can be found in Baraveli *et al.* (2008) and Yu *et al.* (2009).

Two basic characteristics of the planar MOSFET are the threshold voltage V_{th} and the off-state leakage current $I_{off.}$ The former is defined as the voltage above which transistor is on and controls the circuit operation. The latter is the source–drain current leaking through the channel when transistor is off, and primarily controls power consumption and memory timing issues. Both depend exponentially (i.e they are very sensitive) on the gate length (defined by the resist line width or CD), and therefore LER/LWR are expected to critically influence their behavior.

In literature, the influence of LER/LWR on V_{th} and I_{off} has been studied, both theoretically and experimentally. Theoretically, 2D compact models have been devised along with more detailed 3D simulations, including even atomistic phenomena. The experimental studies considered both programmed (intentional) and random LER/LWR. In the following sections, we will give a brief account of the main results of these studies. But before doing that, we will discuss the relationship between LER/LWR of resist lines as measured by lithography metrologists and the deviations of the gate length from the nominal value *along the gate width* (usually called gate length roughness (GLR)). Emphasis will be given to the implications of resist LER/LWR in the gate length (CD) variations (CD uniformity) among different transistors and the role of gate width since, as the last edition of ITRS notes, 'At the 28 nm technology node and beyond, the presence of LWR is becoming the biggest portion of CD variation' (ITRS, 2011, p. 47).

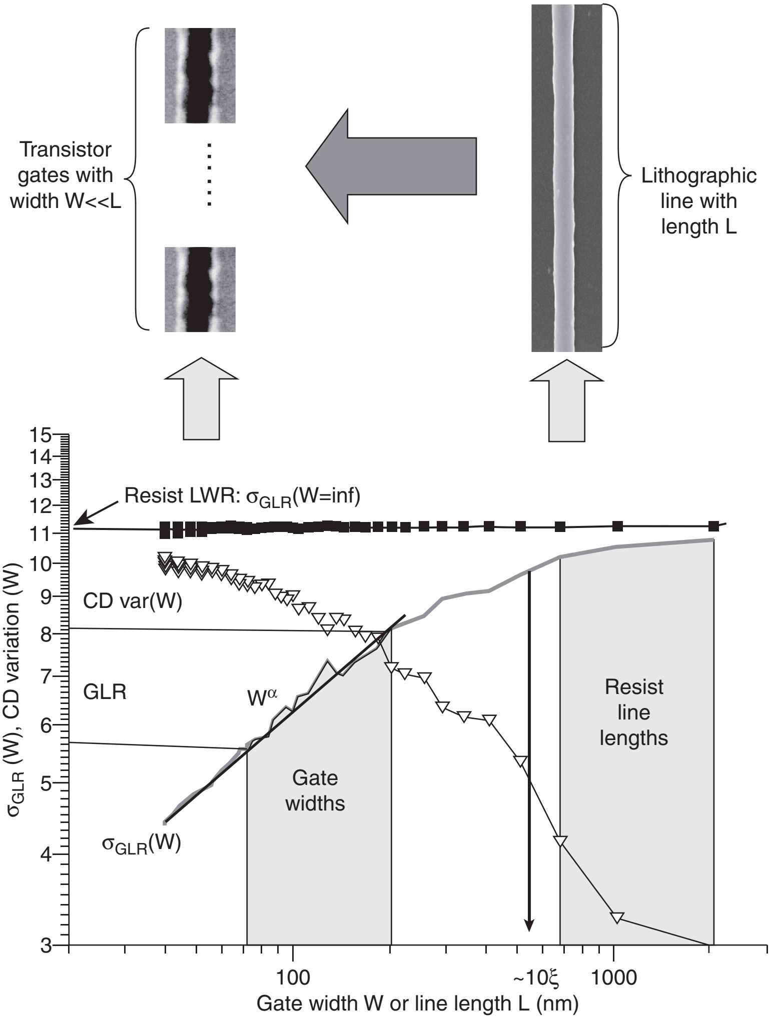
16.5.1 LER/LWR, gate length roughness and CD uniformity

As referred to in Section 16.2.2, resist LWR (rms value) has been agreed to be measured in lines longer than 2 μm (Constantoudis *et al.*, 2004;; Leunissen *et al.*, 2004; Yamaguchi *et al.*, 2005; Patsis *et al.*, 2006; Constantoudis and Gogolides, 2008; Constantoudis *et al.*, 2012.). On the other hand, GLR is defined in lengths equal to the gate widths, which are usually some multiples of CD (~20–40 nm) i.e much shorter than 2 μm. Given the dependence of rms value on edge length (see Section 16.2.2), this length difference translates to differences in the rms values of LWR and GLR. This is shown explicitly in Fig. 16.7, where we can also see that the amount of the rms difference (between LWR and GLR) is determined by the spatial parameters of resist LWR (correlation length ξ and roughness exponent α) and the gate width W. However, this reduction of rms is accompanied by an increase of variation between the average gate length (CD) of transistor gates. Indeed, if we divide a long line into smaller sequential segments, then it can be proved analytically that the sum of the squares of the rms values of the segments and the standard deviation of the CD values remains fixed and equal to the square of the rms of the infinitely long line:

$$\sigma_{\mathrm{LWR}}^2\,(\mathrm{W}) + \mathrm{CDvar}^2(\mathrm{W}) = \sigma_{\mathrm{LWR}}^2\,(\mathrm{W{=}inf})$$

Figure 16.7 graphically illustrates this relationship, as well as its relation to ξ, α. Summarizing, one can say that the resist LWR estimated at long lines is shared in transistor gates into GLR (intra-transistor roughness) and CD variations (inter-transistor roughness), the sharing being controlled by ξ, α. In frequency terms, this means that the low frequency resist LWR contributes to gate CD non-uniformity, while the high frequency fluctuations are associated with GLR.

The message of the above finding is that both GLR and CD variation should be considered to provide a reliable picture of the LWR effects on device performance. Furthermore, it poses the question of the relevant importance of intra and inter-transistor roughness or similarly on the role of ξ, α (besides rms) on transistor yield.



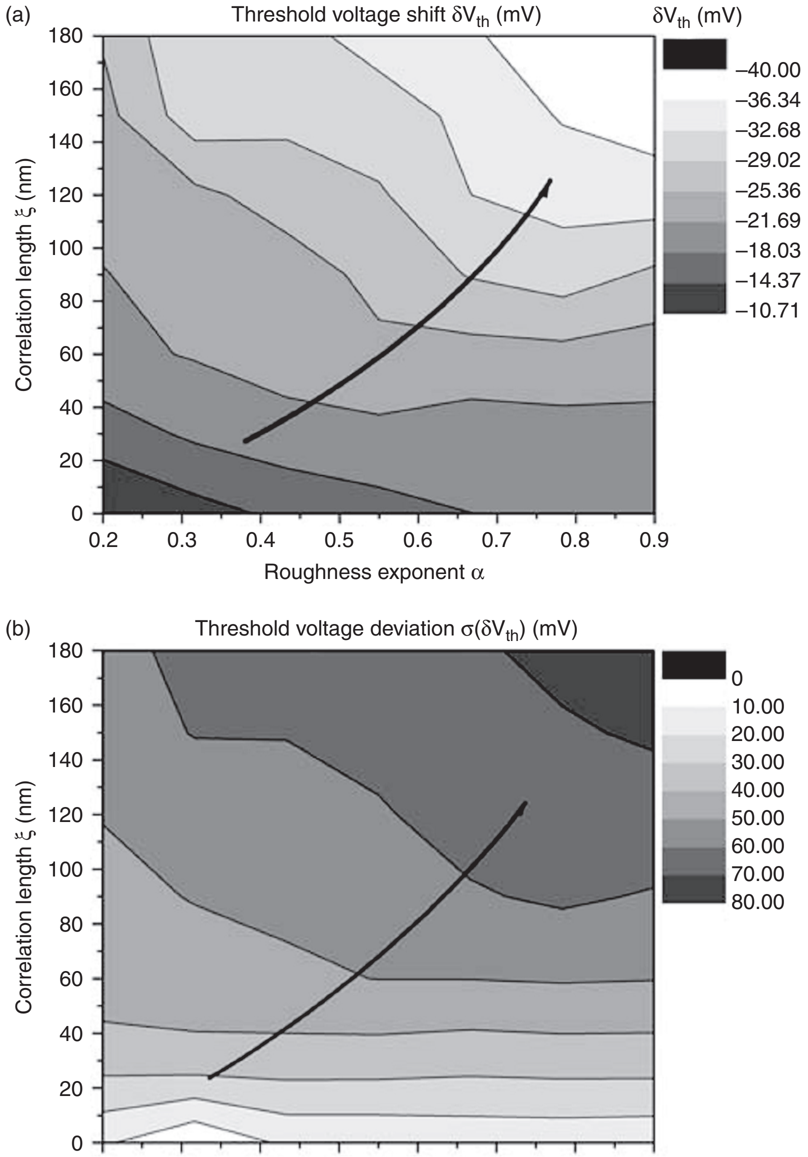
16.7 Rms of GLR (σ_{LWR}) and CD variation (standard deviation) vs the gate width or line length. One can notice the sharing of resist LWR of long lines into GLR and CD variation as well as the critical role of the spatial LWR parameters (ξ, α) in the control of this partition (Constantoudis and Gogolides, 2008).

16.5.2 LER effects on transistor performance

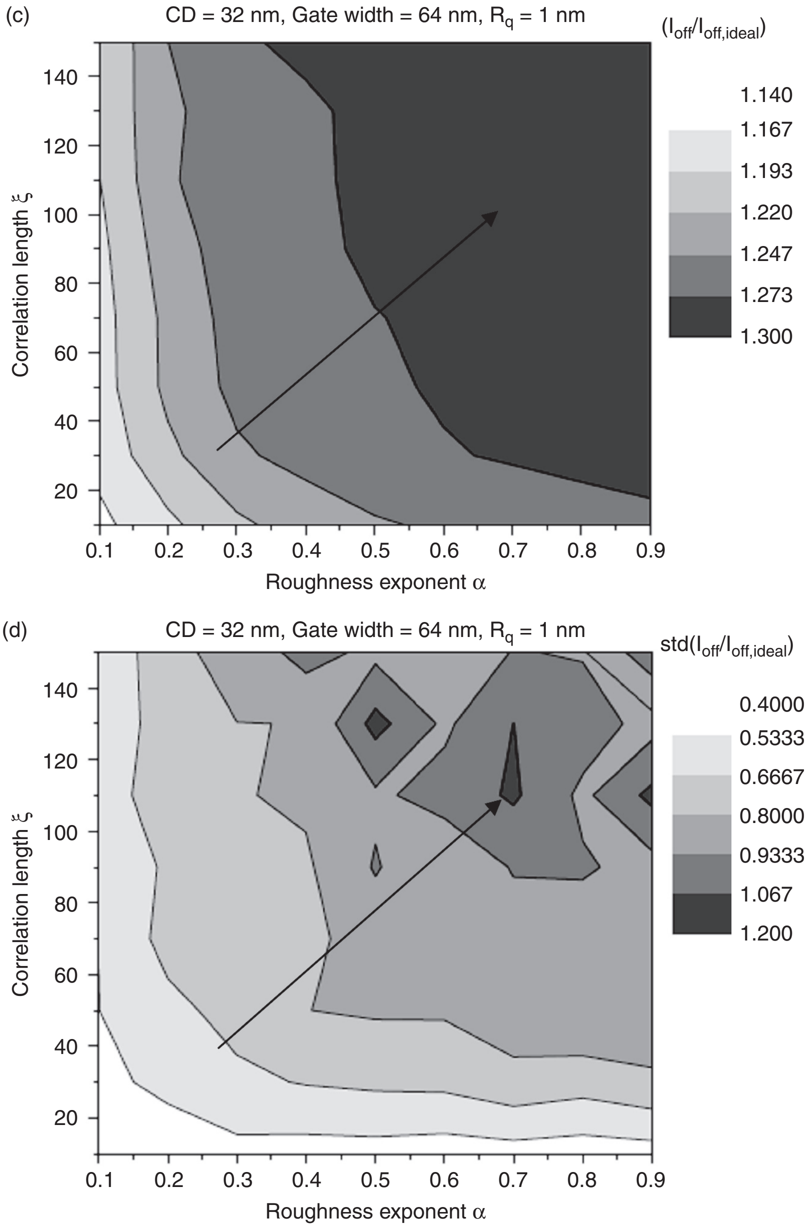
Modeling and simulation

Modeling LER effects on transistor performance is simplified by dividing each transistor into subtransistors, which are assumed to have a rectangular shape with no roughness. (Oldiges *et al.*, 2000; Diaz *et al.*, 2001; Croon *et al.*, 2003; Yamaguchi *et al.*, 2003a; Kim *et al.*, 2004; Yamaguchi *et al.*, 2005; Lorusso *et al.*, 2006; Patsis *et al.*, 2006; Constantoudis *et al.*, 2007, 2008). This method was proposed first by Oldiges *et al.* (2000), and was used to obtain an analytical model and a first comparison with experimental results by Diaz *et al*. The length of each subtransistor varies, and equals the average of the few linewidths (gate lengths) within it. Thus, the whole transistor is represented as a union of rectangular subtransistors with the same widths but varying lengths which represent the GLR of the transistor. Due to the rectangular shape of each subtransistor, one can apply simple analytical formulas incorporating short channel effects for the estimation of the drain and the off-state leakage current I_{off} for each subtransistor. The threshold voltage of the transistor can be estimated through a threshold value $I_{\text{ds,thres}}$ from the total $I_{\text{ds}}(V)$ curve calculated from the summation over all subtransistor currents. Initial applications of this methodology provided the first quantifications of the degradation effects of GLR on both average and standard deviations of V_{th} and I_{off} and made rough predictions for the allowable limit values. Also, with a slight but critical modification of the initial input, this modeling approach has been used to examine the relative importance of intra- and inter-transistor roughness. Instead of generating model edges for the transistor gates with fixed CD and various GLR, first a long fractal self-affine line can be produced modeling the resist line, and then sequential segments of this line can be taken to represent the transistor gates in the above modeling methodology. In this way the total effect of resist LWR at transistor level (both GLR and gate CD variation) can be considered, while equivalently the role of resist spatial LWR parameters ξ, α is illustrated. Figure 16.8 shows the impact of resist ξ, α on mean and standard deviation of V_{th} (*a*, *b*) and $I_{\text{off}}/I_{\text{off,ideal}}$ (*c*, *d*). In all cases, smaller deviations from the ideal (no roughness) behavior are observed at low ξ, α values which correspond to large intra-transistor roughness (GLR) and small CD variation (inter-transistor roughness). This means that if we are able to control the sharing of resist LWR to GLR and CD variation, it is preferable to reinforce GLR against CD variation. Furthermore, this result demonstrates that the degradation effects of resist LWR on transistor electrical performance is mainly due to the inter-transistor CD variation it induces. In the following sub-section we will see that more detailed 3D simulations have reached almost the same conclusion.

In the more advanced and systematic 3D simulations of the effects of LER on conventional MOSFET performance, the Poisson and the current continuity equations are solved self-consistently in the drift–diffusion approximation (Asenov *et al.*, 2003; Roy *et al.*, 2006; Reid *et al.*, 2010a, 2010b). Furthermore, they employ the density gradient (DG) quantum corrections for both electrons and holes to accurately capture quantum confinement effects. LER is introduced into the simulations using a method based on 1D



16.8 Dependence of the average threshold voltage shift δV_{th} (a), the standard deviation $\sigma(\delta V_{\mathrm{th}})$ (b), the average $I_{\mathrm{off}}/I_{\mathrm{off,ideal}}$ (c) and the standard deviation $\sigma(I_{\mathrm{off}}/I_{\mathrm{off,ideal}})$ on the spatial LWR parameters α, ξ. The arrows show the directions for increasing values of all quantities. One can notice that in all cases the modeling predicts that high ξ, α (more inter-transistor roughness and CD variability) result in less reliable transistors (Constantoudis and Gogolides, 2008, 2012.)



16.8 Continued

Fourier synthesis that generates lines using a Gaussian PS. These lines can be considered self-affine fractal with $\alpha = 1$.

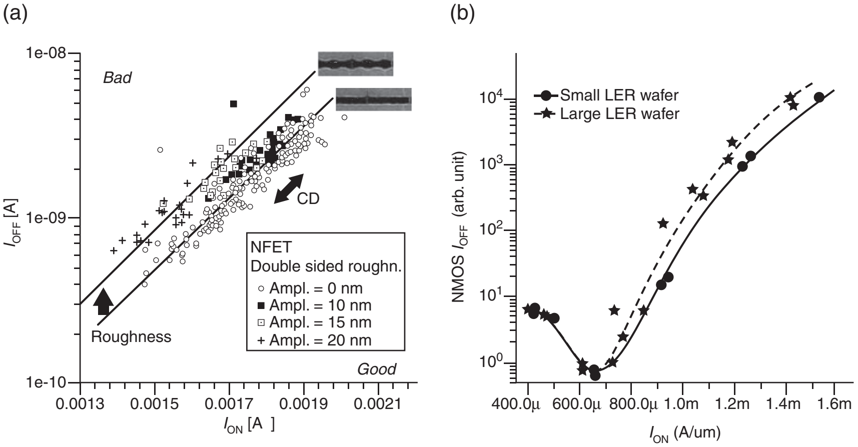
Simulation results for a huge number of sample devices (~10^4–10^5) have shown that LER introduces statistical variability of a comparable magnitude to random discrete dopants (RDD), with the standard deviation being ~40% of the RDD-induced standard deviation. Although RDD remains the dominant source of statistical variability in conventional MOSFETs, a failure to reduce LER below its current level ($3\sigma_{\text{LWR}}$ ~ 5 nm) could promote it as the primary variability source when gate lengths are scaled down to 20 nm and below. Furthermore, a strong nonlinear correlation has been observed between the threshold voltage and the average channel length (CD) of the simulated devices, which governs the LER-induced distribution of V_{th}. Finally, it has been shown that 3D simulations of devices with smooth gate edges (no GLR) but with varying channel length (with CD variation) reproduce very closely the correlation obtained by the full 3D statistical simulations. This result is in agreement with the 2D modeling predictions discussed above, and emphasizes that in short devices resist LER/LWR contributes to statistical variability and degradation mainly through the dimensional variability it induces (see Fig. 16.8).

Experimental results

In order to achieve better isolation and control of LER/LWR effects on device performance, gates with edges having programmed (or intentional) roughness have been fabricated either by e-beam lithography or by exposing a contact layer mask on top of the gate to create local CD variations and hence LER/LWR (Croon *et al.*, 2003; Chandhok *et al.*, 2007; Choi *et al.*, 2008). Due to the large widths of the measured gates and the periodicity of the induced edge roughness (sinusoidal), the full spectrum of LWR was included inside the gates, i.e. no CD variation has been considered. GLR has been controlled by varying the amplitude and period of the regular protrusions, while left and right edges with varying phase shifts have been examined and electrically characterized. Figure 16.9a shows a typical behavior of I_{off} measured for e-beam printed edges with a phase shift 180° for various amplitudes and CDs measured by CD-SEM in etched gates (Choi *et al.*, 2008). Apparently, the increase of amplitude (GLR) results in higher current leakages I_{off} for all I_{on} as well as larger variances.

Experimental results

In a few papers, wafers with resist lines having different LER/LWR have been prepared by applying the optimum 193 nm lithography conditions and then deliberately deviating from these. Then device wafers were fabricated



16.9 I_{on}/I_{off} distributions for (a) programmed 180° shifted edge roughness gates as a function of amplitude (Choi *et al.*, 2008) and (b) random gate edges for two wafers with small and large LER/LWR (Xiong *et al.*, 2004). Both distributions demonstrate the effects of roughness on leakage currents especially at high ion corresponding to gates with smaller CD. Notice the logarithmic scale in y-axis in both diagrams.

through exactly the same process flow, so that transistor with different GLR were obtained (Diaz *et al.*, 2001; Kim *et al.*, 2004; Shibata *et al.*, 2004; Xiong *et al.*, 2004; Gustin *et al.*, 2008). Similar to the programmed LER/LWR studies, long gate widths are used and thus full spectrum LWR is transferred to GLR. An example of measured I_{off} *vs* I_{on} curves for two wafers with small ($1\sigma_{LWR}$ ~ 2 nm) and large ($1\sigma_{LWR}$ ~ 7 nm) is shown in Fig. 16.9b (Xiong *et al.*, 2004). One can notice up to three-fold increase of the I_{off} for gates with high performance (~40 nm nominal gate length).

16.6 Conclusions

The aim of this chapter was to provide the interested reader a short introduction to LER/LWR concepts, a report of the basic findings of the recent intensive research, and an overview of the open challenges, which may shape future research in this field. The emphasis was not on detailed description of the issues, but on the interdisciplinarity of LER/LWR studies and the subsequent and urgent need for transfer of information and research results among various research fields (metrology/statistics, lithography/etching, device physics, and modeling). For example, it may be critical for the LER/LWR metrologists to be aware of which aspect of LER/LWR is more detrimental to device performance, and thus worth being measured, as well as for resist lithographers to know the limits of LER/LWR metrology and measurements so that they employ reliable metrics for the evaluation of the used materials and processes.

Having in mind this demand for enhancing communication among various LER/LWR research groups, we summarize below very briefly the main findings of the recent research as well as the open issues that may play a crucial role in the future of LER/LWR studies.

16.6.1 Metrology: characterization

Advances up to now: Methods exist for noise-free LER/LWR measurements through CD-SEM. First investigations for CD-AFM and scatterometry measurements of LER/LWR have appeared. More detailed (full spectrum) characterization of LER/LWR (three-parameter model) is available.

Challenges: Further advancement of noise-free LER/LWR methods for easier implementation is needed. Comparison of different measurement techniques and calibrations are necessary, as well as the development of a virtual hybrid-holistic metrological approach.

16.6.2 Origins and mitigation of LWR

Advances up to now: RLS triangle and attempts to overcome it (PAG-bound polymers, acid amplifiers, etc.) are now standard in the literature, especially regarding EUV lithography. Post-processing treatments have been shown to mitigate LER efficiently. Both analytical (macroscopic) and numerical (mesoscopic) modeling of LER formation have been developed, capturing the main dependencies of LER on process and material factors.

Challenges: Further reduction of LER/LWR in EUV resists is needed by improving the above-mentioned strategies or new approaches. Further development of mesoscale simulation is needed to describe the full frequency spectrum of LER.

16.6.3 Device effects

Advances up to now: resist LWR effects on gate CD variation are understood especially regarding the dominant role of CD variation aspect of LER/LWR (demonstrated by 2d and 3d simulations and experimental results); atomistic simulations have shown that LER/LWR device effects gain increasing importance as CD shrinks (<20 nm). Furthermore, they have predicted the V_{th} distributions and their relation to gate CD distributions.

Challenges: Enhancement of the statistics of the simulations to include LER variability effects, and comparison of 2D modeling and 3D simulations to devise compact model for LER device effects are needed.

16.7 References

Anderson C N, Naulleau P P, Niakoula D, Hassanein E, Brainard R, Gallatin G and Dean K (2008), 'Influence of base and photoacid generator on deprotection blur in extreme ultraviolet photoresists and some thoughts on shot noise', *J Vac Sci Technol B*, **26**, 2295–2299.

Asenov A, Kaya S and Brown A R (2003), 'Intrinsic parameter fluctuations in decananometre MOSFETs introduced by gate line edge roughness', *IEEE Trans Electron Devices*, **50**, 1254–1260.

Azarnouche L, Pargon E, Menguelti K, Fouchier M, Fuard D, Gouraud C, Verove C and Joubert O (2011),'Unbiased line width roughness measurements with critical dimension scanning electron microscopy and critical dimension atomic force microscopy', *J Appl Phys*, **111**, 084318.

Baraveli E, Jurczak M, Speciale N, De Meyer Kristin and Dixit A (2008), 'Impact of LER and random dopant fluctuations on FinFET matching performance', *IEEE Trans Nanotechnol*, **7**, 291–298.

Biafore J J, Smith M D, van Setten E, Wallow T, Naulleau P, Blankenship D, Robertson and S A, Deng Y (2010), 'Resist pattern prediction at EUV', *Proc SPIE,* **7636**, 76360R1–76360R10.

Brainard R L, Trefonas P, Lammers J H, Cutler C A, Mackevich J F, Trefonas A and Robertson S A (2004), 'Shot noise, LER, and quantum efficiency of EUV photoresists', *Proc SPIE*, **5374**, 74–85.

Brill B, Gov S, Hak D, Sorin V, Marcu T and Bunday B (2010), 'LER detection using dark field spectroscopic reflectometry', *Proc SPIE*, **7638**, 76380P.

Bunday B D, Bishop M, McCormack D, Villarrubia J S, Vladar A E, Dixson R, Vorburger T and Orji N G (2004), 'Determination of optimal parameters for CD-SEM measurement of line edge roughness', *Proc SPIE*, **5375**, 515–533.

Chandhok M, Datta M, Lionberger D and Vesecky S (2007), 'Impact of line width roughness on Intel's 65 nm process devices', *Proc SPIE*, **6519**, 65191A.

Chandhok *M*,Frasure *K*, Putna E S, Younkin T R, Rachmady W, Shah U and Yueh W (2008), 'Improvement in linewidth roughness by postprocessing', *J Vac Sci Technol B*, **26**, 2265–2270.

Choi K -H, Dittrich R, Goldbach M, Hohle C, Keil K, Marschner T, Tesauro M, Thrum F, Zimmermann R and Kretz J (2008), 'Gate edge roughness in electron beam direct write and its influence to device characteristics', *Proc SPIE*, **6921**, 69210J.

Constantoudis V, Patsis G. and Gogolides E (2007), 'Correlation length and the problem of Line Width Roughness', *Proc. SPIE*, **6518**, 65181N.

Constantoudis V, Patsis G P, Tserepi A and Gogolides E (2003), 'Quantification of line-edge roughness of photoresists. II. Scaling and fractal analysis and the best roughness descriptors', *J Vac Sci Technol B*, **21**, 1019–1026.

Constantoudis V, Patsis G P and Gogolides E (2004a), 'Photoresist line-edge roughness analysis using scaling concepts', *J Micro/Nanolith MEMS MOEMS*, **3**, 429–436.

Constantoudis V, Patsis G P, Leunissen L H A and Gogolides E (2004b), 'Line edge roughness and critical dimension variation: Fractal characterization and comparison using model functions', *J Vac Sci Technol B*, **22**, 1974–1983.

Constantoudis V and Gogolides E (2008), 'Fractal dimension of line width roughness and its effects on transistor performance', *Proc SPIE,* **6922**, 692223.

Constantoudis, Kokkoris G, Gogolides E, Pargon E and Martin M (2010), 'Effects of resist sidewall morphology on line-edge roughness reduction and transfer during etching: Is the resist sidewall after development isotropic or anisotropic?', *J Micro/Nanolith MEMS MOEMS*, **9**, 041209.

Constantoudis V, Patsis G P and Gogolides E (2012), 'Fractals and device performance variability: The key role of roughness in micro and nanofabrication', *Microelectron Eng*, **90**, 121–125.

Croon J A, Leunissen L H A, Jurczak M, Benndorf M, Rooyackers R, Ronse K, Decoutere S, Sansen W and Maes H E (2003), 'Experimental investigation of the impact of line-edge roughness on MOSFET performance and yield', *Proc ESSDERC*, 227–230.

Diaz C H, Tao H -J, Ku Y -C, Yen A and Young K (2001), 'An experimentally validated analytical model for gate line edge roughness', *IEEE Elect Dev Lett*, **22**, 287–289.

Fedynyshyn, T H, Astolfi D K, Cabral A, Cann S, Pottebaum I and Roberts J M (2009), 'PAG segregation during exposure affecting innate material roughness', *Proc SPIE,* **7273**, 727349.

Foldyna M, Germer T A, Bergner B C and Dixson R G (2011), 'Generalized ellipsometry of artificially designed line width roughness', *Thin Solid Films,* **519**, 2633–2636.

Foubert R, Pret A V, Sanchez E A and Gronheid R (2011), 'Impact of post-litho LWR smoothing processes on the post-etch patterning result', *Proc SPIE*, **7972**, 797213.

Foucher J, Pikon A Andes C and Thackeray J (2007), 'Impact of acid diffusion length on resist LER and LWR measured by CD-AFM and CD SEM', *Proc SPIE*, **6518**, 65181Q.

Gallatin G M (2005), 'Resist blur and line edge roughness', *Proc SPIE*, **5754**, 38–52.

Gallatin G M, Naulleau P P, Niakoula D, Brainard R, Hassanein E, Matyi R, Thackeray J, Spear K and Dean K (2008), 'Resolution, LER and sensitivity limitations of photoresist', *Proc SPIE*, **6921**, 69211E.

Gallatin G M and Naulleau P P (2011), 'Modeling the transfer of line edge roughness from an EUV mask to the wafer', *Proc SPIE*, **7969**, 796903.

Garidis K, Pret A V and Gronheid R (2012), 'Mask roughness impact on extreme UV and 193 nm immersion lithography', *Microelectron Eng*, **98**, 138–141.

George S A, Naulleau P P, Krishnamoorthy A, Wu Z, Rutter E W, Kennedy J T, Xie S Y, Flanigan K Y and Wallow T I (2010), 'Characterization of line-edge roughness (LER) propagation from resists: Underlayer interfaces in ultrathin resist films', *Proc SPIE* **7636**, 763605.

Goldfarb D L, Mahorowala A P, Gallatin G, Petrillo K E, Temple K, Angelopoulos M, Rasgon S, Sawin H H, Allen S D, Lawson M C and Kwong R W (2004), 'Effect of thin-film imaging on line edge roughness transfer to underlayers during etch processes', *J Vac Sci Technol B*, **22**, 647–653.

Gronheid R, Rathsack B, Bernard S, Pret A V, Nafus K and Hatakeyama S (2008), 'Effect of PAG distribution on ArF and EUV resist performance', *J Photopol Sci Technol*, **22**, 97–104.

Guo W and Sawin H (2009), 'Review of profile and roughening simulation in microelectronics plasma etching', *J Phys D Appl Phys,* **42**, 194014.

Gustin C, Leunissen L, Mercha A, Decoutere S and Lorusso G (2008), 'Impact of line width roughness on the matching performances of next-generation devices', *Thin Solid Films*, **516**, 3690–3696.

He, D and Cerrina F (1998), 'Process dependence of roughness in a positive-tone chemically amplified resist', *J Vac Sci Technol B,* **16**, 3748–3751.

Higgins H, Settens C, Wolfe P, Petrillo K, Auger R, Matyi R and Brainard R (2011), 'Coefficient of thermal expansion (CTE) in EUV lithography: LER and adhesion improvement', *Proc SPIE*, **7972**, 797211.

Hiraiwa A and Nishida A (2009), 'Discrete power spectrum of line width roughness', *J Appl Phys*, **106**, 074905.

International Technology Roadmap for Semiconductors (ITRS) (2007), Semiconductor Industry Association.

International Technology Roadmap for Semiconductors (ITRS) (2011), Semiconductor Industry Association.

Kim H W, Lee J Y, Shin J, Woo S -G, Cho H -K and Moon J -T (2004), 'Experimental investigation of the impact of LWR on sub-100-nm device performance', *IEEE Trans Electron Devices,* **51**, 1984–1988.

Kokkoris G, Constantoudis V and Gogolides E (2009), 'Nanoscale roughness effects at the interface of lithography and plasma etching: Modeling of line-edge-roughness transfer during plasma etching' *IEEE Trans. Plasma Sci*, **37**, 1705–1714.

Kudo T, Chakrapani S, Lin G, Anyadiegwu C, Antonio C, Parthasarathy D, Dammel R R and Padmanaban M (2007), 'Some non-resist component contributions to LER and LWR in 193 nm lithography', *Proc SPIE*, **6519**, 651941.

Lawrence W G (2004), 'Spatial frequency analysis of line edge roughness in nine chemically related photoresists', *Proc SPIE*, **5039**, 713–811.

Lawson R A and Henderson C L (2010a), 'Understanding the relationship between true and measured resist feature critical dimension and line edge roughness using a detailed scanning electron microscopy simulator', *J Vac Sci Technol B*, **28**, C6H34–C6H39.

Lawson R A and Henderson C L (2010b), 'Three-Dimensional mesoscale model for the simulation of LER in photoresists', *Proc SPIE*, **7639**, 76392G01–76392G10.

Lee C T, Wang M, Jarnagin N D, Gonsalves K E, Roberts J M, Yueh W and Henderson C L (2007), *Proc. SPIE*, **6519**, 65191E.

Leunissen L H A, Lawrence W G and Ercken M (2004), 'Line-edge roughness: Experimental results related to a two-parameter model', *Microelectron Eng,* **73–74**, 265–270.

Leunissen L H A, Lorusso G F, Ercken M, Croon J A, Yang H, Azordegan A and DiBiase T (2005), 'Full spectral analysis of line width roughness', *Proc. SPIE*, **5752**, 578–590.

Li Y G, Mao S F, Li H M, Xiao S M and Ding Z J (2008), 'Monte Carlo simulation study of scanning electron microscopy images of rough surfaces', *J Appl Phys*, **104**, 064901.

Lorusso G F, Leunissen L H A, Gustin C, Mercha A, Jurczak M, Marchman H M and Azordegan A (2006), 'Impact of line width roughness on device performance', *Proc SPIE,* **6152** I, 61520W.

Mack C A (1987), 'Development of positive resists', *J Electrochem Soc,* **134**, 148–152.

Mack C (2009), 'Impact of mask roughness on wafer line-edge roughness', *Proc SPIE*, **7488**, 748828.

Mack C A (2010), 'Line-edge roughness and the ultimate limits of lithography', *Proc SPIE*, **7639**, 763931.

Martin Y and Kumar Wickramasinghe H (1994), 'This should come alphabetically after Mandelbrot, abov Method for imaging sidewalls by atomic force microscopy', *Appl Phys Lett,* **64**, 2498–2500.

Mandelbrot B (1982), *The Fractal Geometry of Nature*, New York, W.H. Freeman and company.

Michaelson T B, Pawloski A R, Acheta A, Nishimura Y and Willson C G (2005), 'The effects of chemical gradients and photoresist composition on lithographically generated line edge roughness', *Proc SPIE*, **5753**, 368–379.

Morita H and Doi M (2010), 'Meso-scale simulation of the line-edge structure based on polymer chains in the developing and rinse process', *Proc SPIE*, **7639**, 763932.

Namatsu H, Nagase M, Yamaguchi T, Yamazaki K and Kurihara K (1998), 'Influence of edge roughness in resist patterns on etched patterns', *J Vac Sci Technol B*, **16**, 3315–3323.

Naulleau P P (2009), 'Correlation method for the measure of mask-induced line-edge roughness in extreme ultraviolet lithography', *Appl. Optics*, **48**, 3302–3307.

Naulleau P P, George S A and McClinton B M (2010), 'Mask roughness and its implications for LER at the 22- and 16-nm nodes', *Proc SPIE*, **7636**, 76362H.

Oehrlein G S, Phaneuf R J and Graves D B (2010), 'Plasma-polymer interactions: A review of progress in understanding polymer resist mask durability during plasma etching for nanoscale fabrication', *J Vac Sci Technol B*, **29**, 010801.

Oldiges P, Lin Q, Petrillot K, Sanchez M, Ieong M and Hargrove M (2000), 'Modeling line edge roughness effects in sub 100 nanometer gate length devices', *International Conference on Simulation of Semiconductor Processes and Devices SISPAD*, 6–8 September, Sheraton Seattle Hotel and Towers, Seattle, WA, USA, 131–134.

Orji N G, Vorburger T V, Fu J, Dixson R G, Nguyen C V and Raja J (2005), 'Line edge roughness metrology using atomic force microscopes', *Meas Sci Technol*, **16**, 2147–2154.

Pargon E, Martin M, Thiault J, Joubert O, Foucher J and Lill T (2008), 'Linewidth roughness transfer measured by critical dimension atomic force microscopy during plasma patterning of polysilicon gate transistors', *J Vac Sci Technol B*, **26**, 1011–1020.

Pargon E, Martin M, Menguelti K, Azarnouche L, Foucher J and Joubert O (2009), 'Plasma impact on 193 nm photoresist linewidth roughness: Role of plasma vacuum ultraviolet light', *Appl Phys Lett,* **94**, 103111.

Patel K, Lahiri S N and Spanos C J (2010), 'Robust estimation of line width roughness parameters', *J Vac Sci Technol B*, **28**, C6H18.

Patsis G P and Gogolides E (2001), 'Simulation of surface and line-edge roughness formation in resists', *Microelectron Eng*, **57–58**, 563–569.

Patsis G P, Constantoudis V, Tserepi A, Gogolides E and Grozev G (2003), 'Quantification of line-edge roughness of photoresists. I. A comparison between off-line and on-line analysis of top-down scanning electron microscopy images', *J Vac Sci Technol B*, **21**, 1008–1018.

Patsis G P and Gogolides E (2005), 'Material and process effects on line-edge-roughness of photoresists probed with a fast stochastic lithography simulator', *J Vac Sci Technol B*, **23**, 1371–1375.

Patsis G P, Constantoudis V and Gogolides E (2006), 'Integrated simulation of line edge roughness (LER) effects on sub-65 nm transistor operation : From lithography simulation, to LER metrology, to device operation', *Proc SPIE,* **6151** II, 61513J.

Patsis G P, Drygiannakis, Constantoudis V, Raptis and Gogolides E (2010), 'Stochastic modeling and simulation of photoresist surface and line-edge roughness evolution', *Eur Polym J*, **46**, 1988–1999.

Pawloski A R and Nealey P F (2002), 'Effect of photoacid generator concentration on sensitivity, photoacid generation, and deprotection of chemically amplified resists', *J Vac Sci Technol B*, **20**, 2413–2420.

Pawloski A R, Acheta A, Lalovic I, La Fontaine B M and Levinson H J (2004), 'Characterization of line-edge roughness in photoresist using an image fading technique', *Proc SPIE*, **5376**, 414–425.

Pawloski A R, Acheta A, Bell S, La Fontaine B L, Wallow T and Levinson H J (2006), 'The transfer of photoresist LER through Etch', *Proc SPIE,* **6153**, 615318

Philippou A, Mulders T and Scholl E (2007), 'Impact of photoresist composition and polymer chain length on line edge roughness probed with a stochastic simulator', *J Micro/Nanolith MEMS MOEMS*, **6**, 043005.

Pret A V and Gronheid R (2010), 'Line width roughness mitigation in chemically amplified resist by post-litho processes', *Microelectron Eng*, **87**, 1127–1130.

Pret A V and Gronheid R (2011), 'Mask line roughness contribution in EUV lithography', *Microelectron Eng*, **88**, 2167–2170.

Reid D, Millar C, Roy S and Asenov A (2010a), 'Understanding LER-induced MOSFET *VT* variability—Part I: Three-dimensional simulation of large statistical samples', *IEEE Trans. Electron Devices*, **57**, 2801–2807.

Reid D, Millar C, Roy S and Asenov A (2010b), 'Understanding LER-induced MOSFET *VT* variability—Part II: reconstructing the distribution', *IEEE Trans. Electron Devices*, **57**, 2808–2813.

Rice B J, Cao H B and Chandhok M (2003), 'Effects of processing parameters on line-width roughness', *Proc SPIE,* **5039**, 384–392.

Roy G, Brown A R, Adamu-Lema F, Roy R and Asenov A (2006), 'Simulation study of individual and combined sources of intrinsic parameter fluctuations in conventional nano-MOSFETs', *IEEE Trans. Electron Devices*, **53**, 3063–3070.

Saeki A, Kozawa T, Tagawa S, Cao H B, Deng H and Leeson M J (2008), 'Line edge roughness after development in a positive-tone chemically amplified resist of post-optical lithography investigated by Monte Carlo simulation and a dissolution model', *Nanotechnology*, **19**, 0157051–0157055.

Schnattinger T and Erdmann A (2008), 'A comprehensive resist model for the prediction of line-edge roughness material and process dependencies in optical lithography', *Proc SPIE*, **6923**, 69230R1–69230R12.

Schuster T, Rafler S, Ferreras Paz V, Frenner K, and Osten W (2009), 'Fieldstitching with Kirchhoff-boundaries as a model based description for line edge roughness (LER) in scatterometry', *Microelectron Eng* **86**, 1029–1032.

Shibata K, Izumi N and Tsujita K (2004), 'Influence of line edge roughness on MOSFET devices with sub-50 nm gates', *Proc SPIE*, **5375**, 865–875.

Shin J, Han G, Ma Y, Moloni K and Cerrina F (2001), 'Resist line edge roughness and aerial image contrast', *J Vac Sci Technol B*, **19**, 2890–2895.

Shyu D -M, Ku Y -S and Smith N (2007), 'Angular scatterometry for line-width roughness measurement', *Proc SPIE*, **6518**, 65184G.

Smith B (2012), 'Optical projection lithography', in Feldman M, *Nanolithography*, Cambridge, Woodhead.

Soda E, Kondo S, Saito S, Koyama K, Jinnai B and Samukawa S (2009), 'Mechanism of reducing line edge roughness in ArF photoresist by using CF_3I plasma', *J Vac Sci Technol B*, **27**, 2117–2121.

Tsiartas P C, Flanagin L W, Henderson C L, Hinsberg W D, Sanchez I C, Bonnecaze R T and Willson C G (1997), 'The mechanism of phenolic polymer dissolution: A new perspective', *Macromolecules*, **30**, 4656–4664.

Van Steenwinckel D and Lammers J H, Koehler T, Brainard R L and Trefonas P (2006), 'Resist effects at small pitches', *J Vac Sci Technol B*, **24**, 316–320.

Van Steenwinckel D Lammers J H, Leunissen L H A and Kwinten J A J M (2005), 'Lithographic importance of acid diffusion in chemically amplified resists', *Proc SPIE* **5753**, 269–280.

Villarrubia J S (2005), 'Issues in line edge and line width roughness metrology', *Characterization and Metrology for ULSI Technology*, *AIP Conference Proceedings,* Richardson, Texas, 15–18 March, **788**, 386–393.

Villarrubia J S and Bunday B D (2005), 'Unbiased estimation of linewidth roughness', *Proc SPIE,* **5752**, 480–489.

Wang C, Choi K -W, Jones R L, Soles C, Lin E L, Wu W, Clarke J S, Villarrubia J S and Bunday B (2008), 'Linewidth roughness and cross-sectional measurements of sub-50 nm structures with CD-SAXS and CD-SEM', *Proc SPIE*, **6922**, 69221Z.

Xiong S, Bokor J, Xiang Q, Fisher P, Dudley I, Rao P, Wang H and En B (2004), 'Is gate line edge roughness a first-order issue in affecting the performance of deep sub-micro bulk MOSFET devices', *Trans Semi Manuf*, **17**, 357–361.

Yamaguchi A, Tsuchiya R, Fukuda H, Komuro O, Kawada H and Iizumi T (2003a), 'Characterization of line-edge roughness in resist patterns and estimations of its effect on device performance' *Proc SPIE,* **5038**, 689–698.

Yamaguchi A and Komuro O (2003b), 'Characterization of line edge roughness in resist patterns by using Fourier analysis and auto-correlation function', *Jpn J Appl Phys*, **42**, 3763–3770.

Yamaguchi A, Fukuda H, Arai T, Yamamoto J, Hirayama T, Shiono D, Hada H and Onodera J (2005), 'Spectral analysis of line-edge roughness in polyphenol EB-resists and its impact on transistor performance', *J Vac Sci Technol B*, **23**, 2711–2715.

Yamaguchi A, Steffen R, Kawada H and Iizumi T (2006), 'Bias-free measurement of LER/LWR with low damage by CDSEM', *Proc SPIE,* **6152**, 61522D1–61522D8.

Yamaguchi A and Yamamoto J (2008), 'Influence of image processing on line-edge roughness in CD-SEM measurement', *Proc SPIE,* **6922**, 692221.

Yu S, Zhao Y, Du G, Kang J, Han R and Liu X (2009), 'The impact of line edge roughness on the stability of a FinFET SRAM', *Semicond Sci Technol*, **24**, 025005.

17

New applications and emerging technologies in nanolithography

F. YESILKOY, C. ROPP, Z. CUMMINS, R. PROBST,
E. WAKS, B. SHAPIRO and M. PECKERAR,
University of Maryland, USA

DOI: 10.1533/9780857098757.538

Abstract: The infrared (IR) spectrum lies between the microwave and optical frequency ranges, which are well suited for communication and energy harvesting purposes, respectively. The long wavelength IR (LWIR) spectrum, corresponding to wavelengths from 8 µm to 15 µm, includes the thermal radiation emitted by objects at room temperature and the Earth's terrestrial radiation. Therefore, LWIR detectors are very appealing for thermal imaging purposes. In this chapter, we investigate the prospects of Mid-IR antenna coupled Metal-Insulator-Metal rectifying diodes to be used for LWIR detection and harvesting purposes. Considering the research presented in the literature on this subject, we introduce current challenges that lead to the future research directions. Moreover, we support the analysis on the antenna coupled tunneling diodes with our most recent results to draw a solid picture.

Key words: rectenna, energy harvester, IR detector, MIM junction/diode, tunneling, junction/diode resistance, junction/diode capacitance, tunneling oxide/barrier, tunnel junction/diode, current-voltage asymmetry, geometric asymmetry, rectification, EBL, lift-off, parasitic impedance, Antenna Coupled Metal Insulator Metal (ACMIM), Mid-IR, Terahertz, cut-off frequency, RC time constant, Geometric decoupling, Nonlinearity, Sensitivity, Responsivity, Quantum efficiency, Antenna efficiency, Impedance matching, Bow-tie antenna, Finite Element Method, AC harmonic simulations, E-field enhancement, Nickel, IR imaging.

17.1 Introduction

Undoubtedly, the microelectronic industry has been a major driver in the field of high-resolution patterning and pattern placement. For many years, Moore's Law improvements in the functionality of digital technology were driven almost exclusively by improvements in lithography. Now, this is no longer the case. That is not to say that such improvements are unimportant. As described in this volume, advanced design rule scaling is still an active area of scientific and technological research. New techniques, such as 'double patterning,' and new tools, such as multi-beam e-beam pattern generators, still garner a fair amount of interest in the chip-making industry. But developmental emphasis within the 'chip' community is shifting.

This is the result of two factors. In the past, shrinking component dimensions was a sure path to speed and to increased component density. Clearly, this is no longer the case. Over the last decade, we have engaged in exponential reduction in minimum feature size, but maximum clock-speed has remained in the vicinity of 3 GHz. Transistor count has increased by the Moore's Law rule of doubling every two years. But some of this performance-increase stems from the effect of successful fielding of 'multi-core' technologies utilizing a number of parallel processor chips instead of a single large-area chip.

In addition, the component density problem is addressed today by using aggressive materials technologies that augment the lithographic tool set. For example, fully depleted, ultrathin silicon-on-insulator structures aid in density improvement by limiting space–charge encroachment into the field of adjacent devices. Three-dimensional transistors (such as 'finFETs') further increase density by turning the transistor on its side. This pushes a portion of the physical extend of the transistor above the chip surface, rather than running the whole structure parallel to the surface.

To make these chips work, other problems must be surmounted in addition to those presented by patterning. Signals propagate through interconnects that are, in effect, transmission lines. Propagation speed is far less than the speed of light – inversely proportional to the square root of line inductance times the line capacitance. Moving interconnects closer increases inter-line capacitance and mutual inductance. Varying line-density means varying propagation speed, creating 'clock distribution problems.' That is, it is difficult to synchronize different subsystems so that they work together to create a useful computing machine.

Furthermore, geometric scaling leads to current leakage problems that increase power dissipation. Increased part density also leads to increased power dissipation. These problems are at least as difficult to surmount as those encountered in patterning. Thus, given a steady (or decreasing) research budget, patterning technology funding must surrender at least part of its resource to surmounting these other barriers.

But this does not mean that patterning research is at all receding, or that interest in it is abating. New and more varied applications are rapidly emerging. As our patterning and placement skills have touched on features the size of atoms, we can exploit whole new sets of phenomena based on the principles of quantum mechanics. Lithography is becoming more important in fields such as energy harvesting, terahertz communications, spectroscopy, and quantum-optics, as well as in the emerging fields of plasmonics and metamaterials.

In this chapter we describe, in some detail, two new applications that entwine non-classical physical principles with advanced patterning techniques. The phenomena exploited are tunneling and quantum confinement.

In our discussion of tunneling, we stress the importance of spatial resolution in creating metal–insulator–metal (MIM) tunnel diodes. Our quantum confinement example is that of the quantum dot. Here we stress pattern placement as a dominant requirement.

It must be stressed that a number of new technologies are in development aimed at component technologies that are not in the realm of the traditional integrated circuit (IC). The ability of certain chemical systems to 'self-assemble' into desired patterns has been exploited for many years.[1,2] While a considerable amount of effort has been devoted to applying these 'radical' processing techniques to IC fabrication, they are generally restricted to highly regular pattern arrays (such as arrays of contacts or vias, or long 'grating' like structures of metal lines). While these applications are important, they create specialized 'niches' of use within the standard fabrication context.

More recent trends include the synthesis of 'deterministic patterning' (more or less synonymous with current day lithographic practice) and 'self-assembly.' For example, block co-polymers can naturally form patterns that might be useful in the lithography of contact-hole arrays. The addition of structural templates provided by conventional patterning has been shown to improve the fidelity and reproducibility of self-assembled patterning. For more on this important, emerging field the reader is referred to some of the key literature in this field.[3–6] These studies show that the enhanced material manipulation capability, obtained by synthesizing self-assembly with deterministic patterning, can yield new insights into the basic science of phase transitions. In the text below, we show how new methods of patterning can yield new classes of devices as well as new science.

We begin with one of the most fundamental quantum phenomena: tunneling. While tunneling structures have been investigated extensively in the past, and a Nobel Prize was awarded to Esaki, Giaever and Josephson for this work, the area remains a fertile field for research. Applications in high-frequency RF detection and infrared energy harvesting abound. Ultra-high-resolution lithography enables the operation of these devices at such high frequencies by allowing for the creation of very small junction area (small parasitic capacitance) tunnel diodes. As a result of the very fast tunneling mechanism and the rapid diode response, when some asymmetry between the diode forward and reverse bias resistance exists, high rectification efficiency can be achieved.

Practically, to use tunneling structures as energy harvesters, they must be 'printed' over large areas – extending over many square meters. This cannot be accomplished by a serial writing technique, such as e-beam lithography (EBL). Emerging lithographic technologies, such as stamping, are good candidates to solve these problems. But many of the tunnel junction processes require layer-to-layer alignment. And as we show, misalignment leads to parasitics that limit response speed and render the tunnel junction incapable of rectification into the infrared. Novel self-alignment techniques must be used to augment advanced patterning.[7]

Physical confinement also leads to unique quantum effects. In particular, the 'particle in a box' problem of elementary quantum mechanics leads to physical systems whose 'state structure' mirrors those of a naturally formed atom leading to discrete energy levels of well-defined separation. While particles of atomic dimensions can be formed using a variety of approaches, and these particles can maintain electrons in well-defined energy states whose separation depends on the geometry of the particle, it is a challenge to place these particles at predetermined positions. This placement is frequently necessary to fully exploit a full range of desirable properties. Placement accuracy is an important part of patterning and it will be discussed using the 'quantum dot' example.

Thus, the examples chosen stress two of the basic aspects of microlithography: resolution (in the case of the tunnel junction) and feature placement (in the case of the quantum dot). Recent technological developments are emphasized in both cases.

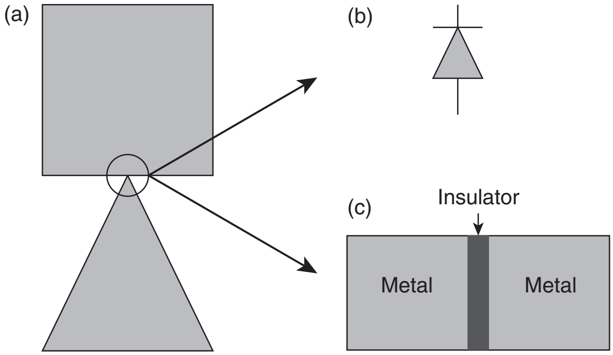
17.2 Applications of high-resolution patterning to new device structures: advances in tunneling structures

For many years, tunneling diodes have held the promise for efficient, high-frequency rectification – rectification well into the terahertz regime. The tunnel process itself is rapid – taking place in timeframes as short as those of atomic transitions (femto-seconds). What limits the response speed is the RC time constant of the tunnel junction. Making small junctions reduces capacitance. But area reduction of the diode increases forward resistance.

An ability to pattern sharply pointed 'field emission' tips in two dimensions improves the situation in two ways. First, a sharp tip concentrates field lines (lightning-rod effect), leading to an increase in tunneling current at a given bias voltage, de-coupling capacitance from forward resistance. Second, diode structure is geometrically asymmetric, leading to asymmetry in the current-voltage response. That is, there is an 'easy' and a 'hard' direction for electron current flow. This asymmetry is essential for rectification to occur at zero bias.

The structure to be realized is shown in Fig. 17.1. On the left (Fig. 17.1a) we see a large square contiguous with the apex of a triangle. The point of contiguity forms a diode (Fig. 17.1b) with an easy electron current flow direction upward from the triangle tip. The material configuration is shown in Fig. 17.1c. A thin insulating layer joins the square and rectangular metal shapes. Insulators normally block current flow. In this case, the insulator is so thin that tunneling can take place across it. This leads to a non-linear current-voltage response curve.

In general, when the MIM junction is geometrically symmetric, the IV curve associated with it is symmetric about the origin. The response is similar to two 'normal' pn-junctions hooked in parallel, yielding forward bias response for either direction of bias polarity. This leads to inefficient



17.1 (a) The basic patterns that comprise the tunnel junction and its associated antenna. (b) The location of the tunnel junction diode. (c) The material configuration of the tunnel diode: MIM.

rectification, as there will be positive and negative current flow when an AC signal is impressed across the diode. This averages to zero net current, delivering no power to load when integrated by a ‘smoothing’ network.

The structure shown, though, concentrates field lines leading to electron emission from the triangle tip when the rectangle is positively biased. When the rectangle is negatively biased, electrons are ‘pushed’ to the triangle tip by a relatively weak electron field, lowering tunnel current significantly. This creates the necessary conduction asymmetry required of a rectifying diode.

Other features of note include the fact that the square and rectangle can be viewed as antennas and their dimensions can be set to receive at specific bands of wavelengths. The configuration shown makes use of both the ‘patch’ and ‘bowtie’ antenna configurations widely used in RF electronics. The resulting integration of the diode and the antenna into a common structure is known as a ‘rectenna.’ This structure has the further advantage of having a near-zero ‘turn-on’ voltage (as contrasted with a pn junction with a 0.7 V turn-on). High turn-on voltage ‘wastes’ incident radiation energy as it leads to no power delivery to a load. As the radiation induced bias drop across a receiving antenna is proportional to the RF power incident, high turn-on voltage also means that incident radiation must be of high energy density to create any response from the rectifying diode. Low light levels are waste in high turn-on voltage diodes.

Thus, it is clear that the geometrically asymmetric tunnel diode structure described has the potential to ‘rectify’ high-frequency radiation. In fact, such devices have been shown to be sensitive to radiation with wavelengths as short as 10 microns. This is the peak of the radiation spectrum created by the cooling earth. The potential application to energy harvesting in the infrared is apparent. But to accomplish this, we must address the following issues:

- The triangular tip should be as sharp as possible
- The tunnel barrier thickness should be on the order of 1 nm to allow reasonable amounts of current to pass
- The tunnel junction area must be small to minimize the introduction of parasitic capacitance.

In the sections below, we deal with each of these issues, in turn. The thrust of our presentation is to show how patterning, process and material technologies must work together to create the desired results.

17.3 Geometry control of the tunnel junctions

The highest resolution achievable on a routine basis is through particle beam technology. Electron beam lithography tools are ubiquitous, and issues in achieving resolution targets have been widely researched. Two areas of electron beam lithography are particularly important: proximity control and alignment.

17.3.1 The future of beam-based lithography: proximity control

Resolution is defined by a number of different physical phenomena in lithography processes. In optical lithography, diffraction ultimately limits resolution. Diffraction effects do not limit resolution in EBL. The wavelength of the electrons shot in the beam is far smaller than target feature sizes. Indeed, EBL systems can focus electrons to spot sizes smaller than 1 nm in diameter when incident on the patterned surface (see Chapter 3). However, as soon as the highly focused electron beam enters the resist, electron interaction in the resist and in the substrate cause undesired dose to be delivered to points that are distant from the incident beam point. This unwanted exposure on unintended regions of the resist is called the ‘proximity effect,’ and is the main source of resolution limitations in EBL. New device applications, whose success relies on the ultra-high resolution promised by EBL, must overcome this effect to achieve nanometric dimension control. Techniques for PE control are also discussed in Chapter 3.

17.3.2 Asymmetric antenna fabrication

Recently interest has arisen in the reduction of conventional microwave antennas to sizes suitable for visible wavelengths.[8] By combining the bow-tie shape with a point diode, one can simultaneously detect and rectify weak optical signals. The architecture of asymmetric bow-tie antenna coupled MIM tunnel junction device needs high accuracy in the nanoscale. As in any antenna design, the dimensional fidelity of the asymmetric bow-tie antenna structure determines the capability of the device. Therefore, the antenna segment dimensions need to be constant while the smallest area and the narrowest gap requirement at the tunnel junction (the triangle and the rectangle patterns get closest) are satisfied. The biggest challenge is to avoid the formation of a bridge between the rectangle and the triangle antenna parts caused by the proximity effect.

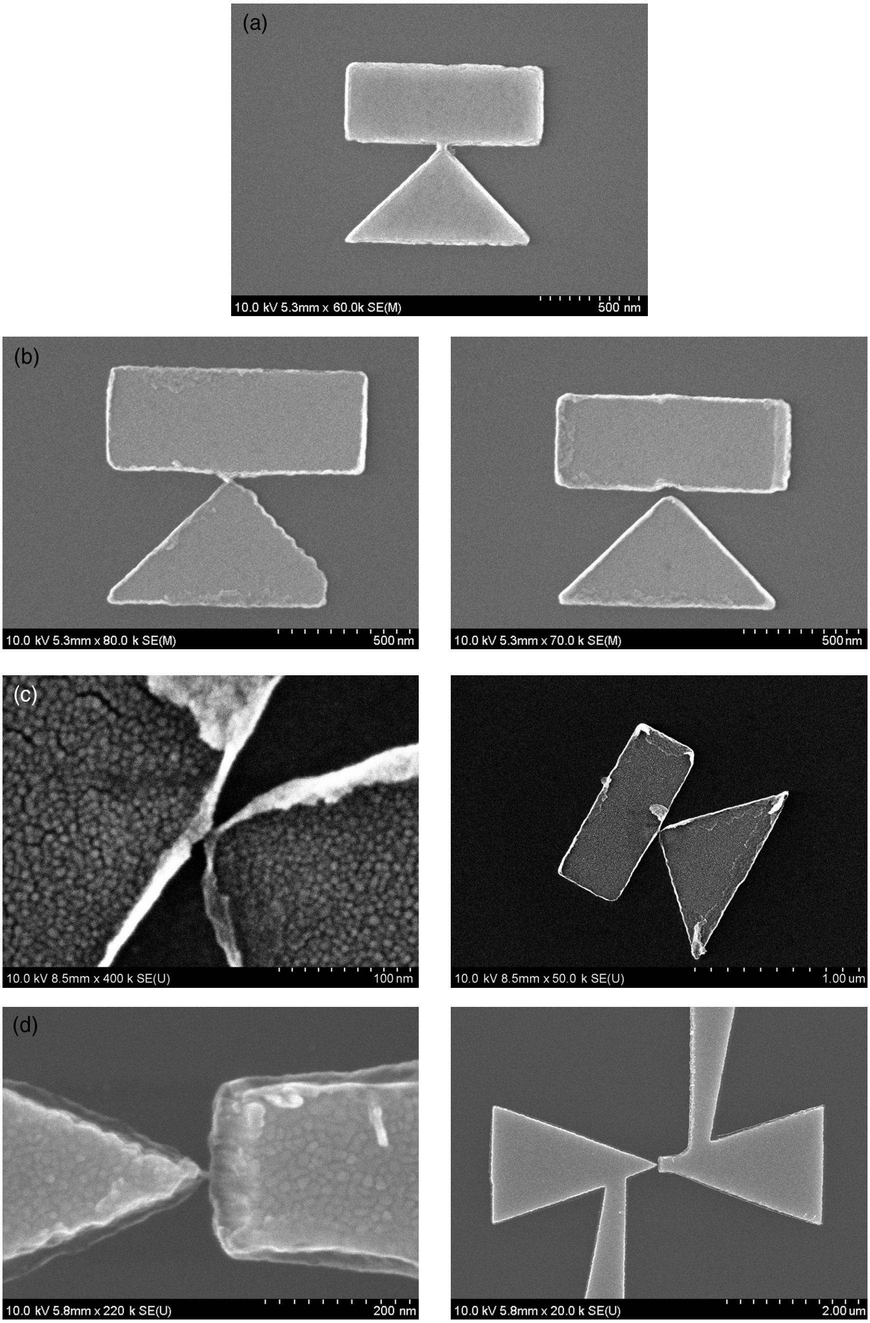
Figure 17.2 shows the results that can be achieved with state-of-the-art processing, with no proximity correction (Fig. 17.2a), conventional proximity correction (Fig. 17.2b), advanced proximity correction[9] (Fig. 17.2c), and two-level patterning[7] (Fig. 17.2d). In addition, a new technique called 'strain assisted lift-off process' has been developed to create complete planar devices. The tunnel diodes fabricated using this technique have extremely small junction areas because the overlapping metal piece detaches due to the strain induced on the top Ni thin film. In Fig. 17.3a and 17.3b, micrographs of an intermediate step where the overlap piece is rolling back and the final planar device are shown, respectively.

In addition to the antenna coupling, surface plasmon enhancement can be employed to maximize the total field at the junction. The most appropriate method of generating surface plasmon is the periodic grating coupling. The period of the grating structure can be determined considering the material parameters. In Fig. 17.4, a micrograph of periodic grating added antenna coupled MIM tunnel diode is shown.

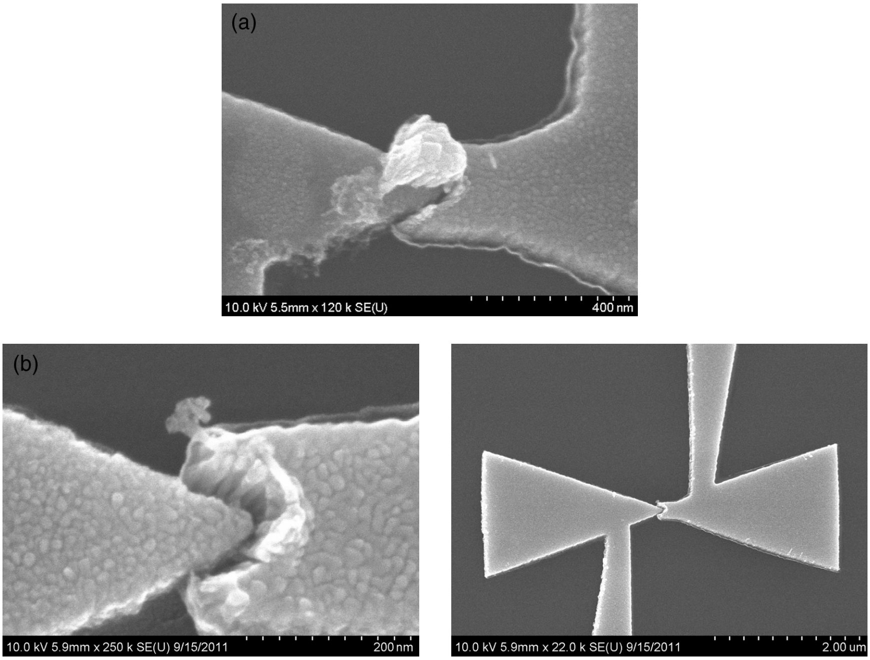
17.3.3 The pattern placement problem

In conventional lithography a projected pattern is aligned to substrate features using alignment marks. These marks are patterns deposited on a substrate providing either optical or beam illumination contrast with the surrounding substrate. We rely on the mask fabrication tool (an e-beam or a laser) to accurately register features within the mask. This is done by moving the exposure beam and the substrate with high precision using electronic steering and interferometer stage motion control. A variety of alignment approaches (many of which are summarized in this volume) are used to tie the position of the mask to the alignment mark. These involve the use of diffraction grains on the substrate and on the mask, moiré pattern approaches and a host of pattern recognition approaches.

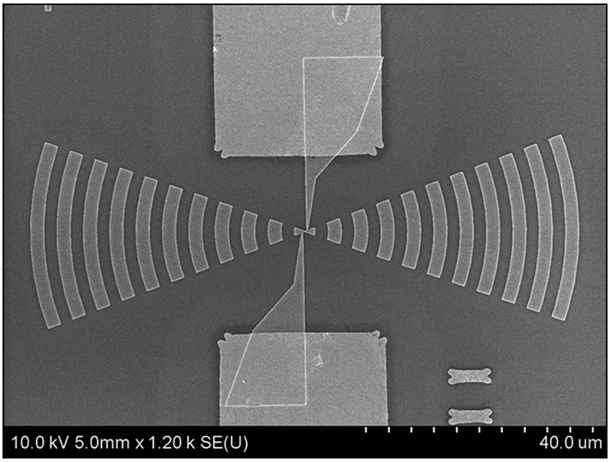
These approaches are aimed at achieving placement in a single system: the rigid mask aligned to marks on a rigid substrate. Some emerging


(a)
10.0 kV 5.3mm x 60.0k SE(M)
500 nm
(b)
10.0 kV 5.3mm x 80.0 k SE(M)
500 nm
10.0 kV 5.3mm x 70.0 k SE(M)
500 nm
(c)
10.0 kV 8.5mm x 400 k SE(U)
100 nm
10.0 kV 8.5mm x 50.0 k SE(U)
1.00 um
(d)
10.0 kV 5.8mm x 220 k SE(U)
200 nm
10.0 kV 5.8mm x 20.0 k SE(U)
2.00 um

17.2 Results achieved on the asymmetric antennas with the state-of-the-art processing. (a) Asymmetric antenna without proximity correction. (b) Conventional proximity correction results. (c) Junction with 2 nm gap achieved on the advanced proximity corrected pattern. (d) two-step lithography technique results.


(a)
10.0 kV 5.5mm x 120 k SE(U)
400 nm
(b)
10.0 kV 5.9mm x 250 k SE(U) 9/15/2011
200nm
10.0 kV 5.9mm x 22.0 k SE(U) 9/15/2011
2.00 um

17.3 Devices fabricated by strain assisted lift-off process. (a) Parasitic capacitance forming overlap piece is rolling back due to the internal strain of the film. (b) Complete planar MIM diode after the overlap piece detached.


10.0 kV 5.0mm x 1.20 k SE(U)
40.0 um

17.4 Periodic grating added antenna coupled MIM tunnel diode.

technologies require more flexibility. For placing objects on chip, there are both contact and non-contact methods. Contact methods refer to using atomic force microscopes (AFM) or micro-/nano-manipulation methods where the object is physically contacted by an AFM tip or by a micro-/nano-arm and is pushed or pulled along the chip surface into place.[10] Since these methods rely on mechanical contact, particle control is sensitive to surface and friction properties.

Non-contact methods such as laser tweezers,[11] dielectrophoretic (DEP) actuation,[12] magnetic[13,14] and flow actuation[15,16] do not directly mechanically contact the particle but instead rely on intermediate forces, such as optical, electrical, magnetic, or viscous drag forces, respectively. Of these, laser tweezers is the most mature and widely used, and it is also the one that has the most parallel capacity – it can manipulate hundreds of particles at once.[17]

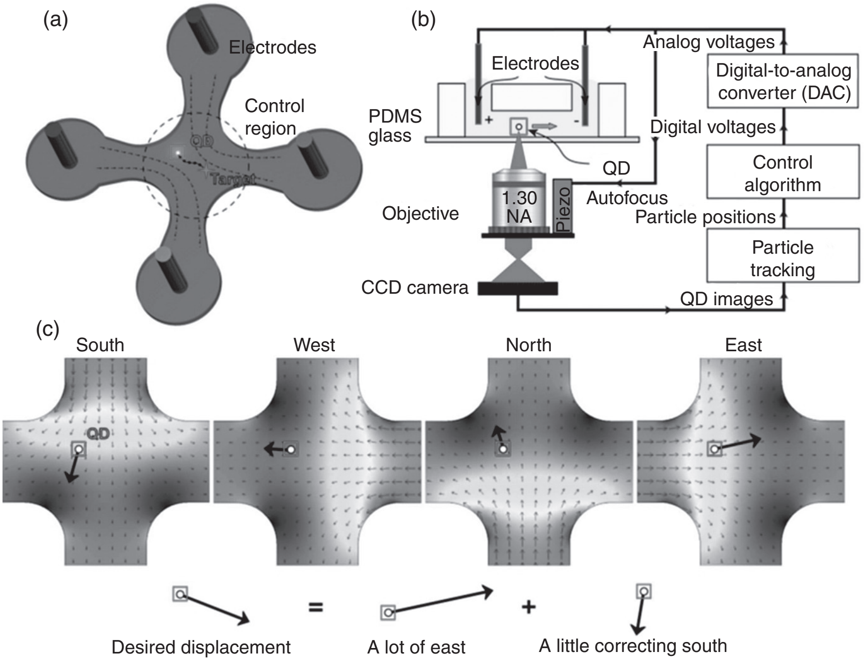
The microfluidic solution

Flow control has two advantages for manipulating nanoscopic particles. The first is that, unlike laser tweezers, DEP, and magnetic actuation, fluid viscous forces vary with particle radius,[18,19] (not with particle volume[20,21]) and this is advantageous for manipulating nanoscopic objects since the viscous forces remain usable. The second is that a moving fluid will apply viscous forces to any object, regardless of its optical or magnetic properties, and so flow control can manipulate particles independent of the optical, electrical, or magnetic properties required for laser tweezing, DEP, or magnetic actuation (see Reference 16).

17.4 The quantum dot placement problem

The positioning of quantum dots is an important example of microfluidic control. The positioning of quantum dots (QDs) in the high field region of nanophotonic devices can enable a strong light–matter interaction for the development of quantum devices.[22,23] In addition, the integration of precisely placed QDs with dielectric nanoparticles and metamaterials provides a promising route towards development of nanoelectronic circuits that can potentially operate at terahertz bandwidths.[24,25]

Nanoscale precision is desired to position the dots within the high electric field regions of photonic and plasmonic structures, typically on the order of ~150 nm in width.[26] This level of accuracy has been demonstrated[15] by placing a cross-channel micro-fluidic device on top of the chip and creating a flow to move each quantum dot from its observed to its desired location. The location of a specific quantum dot is monitored through a microscope in real time as the flow moves the dot to its final intended location.


(a)
Electrodes
Control region
QD
Target
(b)
Analog voltages
Electrodes
Digital-to-analog converter (DAC)
PDMS glass
Digital voltages
QD
Control algorithm
1.30 NA
Piezo
Autofocus
Objective
Particle positions
Particle tracking
CCD camera
QD images
(c)
South
West
North
East
QD
Desired displacement
=
A lot of east
+
A little correcting south

17.5 Experimental setup and control principle. (a) Schematic of the cross-channel microfluidic device structure for 2D control of a quantum dot on a chip. (b) The optical and electronic setup for tracking and feedback control of QDs. (c) Model of the four flow modes resulting from voltages applied to each electrode.[15] (*Source*: Figure taken from Reference 28 and reproduced with permission.)

Two microfluidic channels intersect at a 90° angle (Fig. 17.5). Electroosmotic flow actuation[27] is created by electrodes placed in four fluid reservoirs at the ends of the channels. The colloidal QDs are imaged by an autofocused inverted confocal microscope. After each QD is delivered to the surface of the photonic chip it is fixed in place using a low-viscosity, water-based, negative-tone photoresist.[28,29] A brief ultraviolet laser pulse polymerizes a small cap of fluid immediately around the positioned QD, permanently fixing it on the chip surface.

17.5 Conclusion

Yesterday's emerging technologies have either become mainstream or been forgotten, and it is impossible to predict what will happen next. Nevertheless, it is safe to say that the trend to smaller features will continue, and this chapter explores two ways in which this may happen. For the rest, the best one can do is to look at all the other chapters in this book.

17.6 Acknowledgments

The authors would like to acknowledge the contributions of Prof. Mario Dagenais for his insight into the optical applications of MIM tunnel diodes.

17.7 References

1. Y. S. Lee, *Self-Assembly and Nanotechnology*. John Wiley & Sons, Inc., New York, (2008).
2. B. W. Ninham and P. Lo Nostro, *Molecular Forces and Self Assembly*. Cambridge University Press, Cambridge, (2010).
3. J. K. W. Yang, Y. S. Jung, Chang Jae-Byum, Ross Caroline A. and Berggren Karl K., 'Programmable self-assembly of complex nanoscale patterns,' *Nature Nanotechnology*, vol. **5**, pp. 256–260, (2010).
4. F. J. Castaño, K. Nielsch, C. A. Ross, J. W. A. Robinson and R. Krishnan, 'Anisotropy and magnetotransport in ordered magnetic antidot arrays,' *Applied Physics Letters*, vol. **85**, no. 14, pp. 2872, (2004).
5. H. -S. Kim, L. Bi, H. Paik, D. -J. Yang, Y. C. Park, G. F. Dionne and C. A. Ross, 'Self-assembled single-phase perovskite nanocomposite thin films,' *Nano Letters*, vol. **10**, no. 2, pp. 597–602, (2010).
6. Y. -J. Oh, C. A. Ross, Y. S. Jung, Y. Wang and C. V. Thompson, 'Cobalt nanoparticle arrays made by templated solid-state dewetting,' *Small Weinheim an der Bergstrasse Germany*, vol. **5**, no. 7, pp. 860–865, (2009).
7. F. Yesilkoy, S. Mittal, N. Goldsman, M. Dagenais and M. Peckerar, 'A new process for the fabrication of planar antenna coupled Ni–NiOx–Ni tunnel junction devices,' *Microelectronic Engineering*, vol. **98**, pp, 329–333, (2012).
8. P. Mühlschlegel, H. -J. Eisler, O. J. F. Martin, B. Hecht and D. W. Pohl, 'Resonant optical antennas,'*Science*, vol. **308**, no. 5728, pp. 1607–1609, (2005).
9. F. Yesilkoy, K. Choi, M. Dagenais and M. Peckerar, 'Implementation of e-beam proximity effect correction using linear programming techniques for the fabrication of asymmetric bow-tie antennas,' *SolidState Electronics*, vol. **54**, no. 10, pp. 1211–1215, (2010).
10. C. D. Onal, O. Ozcan and M. Sitti, 'Automated 2-D nanoparticle manipulation using atomic force microscopy,' *Nanotechnology, IEEE Transactions*, vol. **10**, no. 3, pp. 472–481, (2011).
11. D. G. Grier, 'A revolution in optical manipulation,' *Nature*, vol. **424**, no. 6950, pp. 810–816, (2003).
12. C. Zhang, K. Khoshmanesh, A. Mitchell and K. Kalantar-Zadeh, 'Dielectrophoresis for manipulation of micro/nano particles in microfluidic systems,' *Analytical and Bioanalytical Chemistry*, vol. **396**, no. 1, pp. 401–420,(2010).
13. C. Gosse and V. Croquette, 'Magnetic tweezers: micromanipulation and force measurement at the molecular level,' *Biophysical Journal*, vol. **82**, no. 6, pp. 3314–3329, (2002).
14. A. Devries, 'Micro magnetic tweezers for nanomanipulation inside live cells,' *Biophysical Journal*, vol. **88**, no. 3, pp. 2137–2144, (2005).
15. M. D. Armani, S. V. Chaudhary, R. Probst and B. Shapiro, 'Using feedback control of microflows to independently steer multiple particles,' *Journal of Microelectromechanical Systems*, vol. **15**, no. 4, pp. 945–956, (2006).

16. R. Probst, Z. Cummins, C. Ropp, E. Waks and B. Shapiro, 'Flow control of small objects on-chip: manipulating live cells, quantum dots, and nanowires,' *Control System Magazine*, vol. **32, no.** 2, pp. 26–53, (2012).
17. J. Curtis, 'Dynamic holographic optical tweezers,' *Optics Communications*, vol. **207**, no. 1–6, pp. 169–175, (2002).
18. A. Einstein, 'On the movement of small particles suspended in stationary liquids required by the molecular-kinetic theory of heat,' *Annalen der Physik*, vol. **17**, pp. 549–560, (1905).
19. W. Sutherland, 'A dynamical theory of diffusion for non-electrolytes and the molecular mass of albumin,' *Philosophical Magazine*, vol. **9**, no. 54, pp. 781–785, (1905).
20. K. C. Neuman and A. Nagy, 'Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy,' *Nature Methods*, vol. **5**, no. 6, pp. 491–505, (2008).
21. T. B. Jones, 'Basic theory of dielectrophoresis and electrorotation,' *IEEE Engineering in Medicine and Biology Magazine*, vol. **22**, no. 6, pp. 33, (2003).
22. D. E. Chang, A. S. Sorensen, E. A. Demler and M. D. Lukin, 'A single-photon transistor using nanoscale surface plasmons,' *Nature Physics*, vol. **3**, no. 11, pp. 807–812, (2007).
23. A. V. Akimov, A. Mukherjee, C. L. Yu, D. E. Chang, A. S. Zibrov, P. R. Hemmer, H. Park, and M. D. Lukin, 'Generation of single optical plasmons in metallic nanowires coupled to quantum dots,' *Nature*, vol. **450**, no. 7168, pp. 402–406, (2007).
24. N. Engheta, 'Circuits with light at nanoscales: Optical nanocircuits inspired by metamaterials,' *Science*, vol. **317**, no. 5845, pp. 1698–1702, (2007).
25. A. Alu and N. Engheta, 'Tuning the scattering response of optical nanoantennas with nanocircuit loads,' *Natural Photon*, vol. **2**, no. 5, pp. 307–310, (2008).
26. M. Barth, J. Kouba, J. Stingl, B. Löchel and O. Benson, 'Modification of visible spontaneous emission with silicon nitride photonic crystal nanocavities,' *Optics Express*, vol. **15**, no. 25, pp. 17231–17240, (2007).
27. R. F. Probstein, *Physicochemical Hydrodynamics: An Introduction*, 2nd ed. John Wiley & Sons, Inc, (1994).
28. C. Ropp, R. Probst, Z. Cummins, R. Kumar, A. J. Berglund, S. R. Raghavan, E. Waks and B. Shapiro, 'Manipulating quantum dots to nanometer precision by control of flow,' *Nano Letters*, vol. **10**, no. 7, pp. 2525–2530, (2010).
29. C. Ropp, Z. Cummins, R. Probst, S. Qin, J. T. Fourkas, B. Shapiro, and E. Wax, 'Positioning and immobilization of individual quantum dots with nanoscale precision.,' *Nano Letters*, vol. **10**, no. 11, pp. 4673–4679, (2010).

Index

Abbe's model, 451–2
absorbance, 435
absorbance-modulation layer (AML), 187
absorbance-modulation optical lithography (AMOL), 187
accuracy, 456–7
acid diffusion length (ADL), 507, 519
acid hardened resist, 211
acrylate polymerisation, 332
active nanostructures, 403
active optical devices, 381–2
aerial image, 518
'alpha' EUV exposure tools, 51
alternative thermal imprint processes, 327–30
  schematic representation of LADI method for patterning silicon, 329
  SEM image of template and patterned silicon using LADI, 330
amplitude defects, 67
amplitude zone plate, 184
analytical modelling, 514–15
anisotropic etching, 432
anodic bonding, 430–1
anodised aluminium oxide (AAO)
  self-assembly, 361
antenna, 385
  fabrication, 543–4
    devices fabricated by strain assisted lift-off process, 546
    periodic grading added antenna coupled MIM tunnel diode, 546
    results from state-of-the-art processing, 545
antireflective coating (ARC), 65, 181, 225
aperture probe near-field scanning optical microscopy, 390
apertureless probe near-field scanning optical microscopy, 390–1
argon fluoride (ArF) excimer laser, 108, 217
arrayed waveguide gratings (AWG), 381
ASML, 51
aspect ratio (AR), 72
astigmatism, 20
asymmetric systems, 456
atomic force microscopy (AFM), 350, 403, 547
  lithography, 432
attenuated phase shift masks (APSM), 33
autocorrelation function, 510

bar array, 493
beam-based lithography, 543
beam equivalent pressure (BEP), 404
binary masks, 160
binary-phase Fresnel zone plate, 184
'binary' photomask, 31
bio-electrocatalysis, 435
biochemical sensors, 382
biomolecular separations, 436
Bit patterned media (BPM), 334
Blackman-Harris window function, 477
block copolymer lithography, 291–4
  domain structures and their orientation for forming line-and-space and dots/holes, 293
  schematic illustration of process taking vertically oriented cylindrical microdomain, 292

block copolymers, 287–94
  directed self-assembly, 294–302
    schematic diagram of graphoepitaxy, 296
    schematic illustration of chemical heteroepitaxy, 297
    schematic illustration of chemical heteroepitaxy of PS-b-PMMA with PS-brush surface, 299
    SEM image of cylindrical and lamellar microdomain structures by PS-b-PMMA, 295
    SEM image of cylindrical microdomain structure of PS-b-PMMA and 2D-FFT image, 298
    SEM image of EB resist and cylinder structures of PS-b-PMMA, 301
    SEM image of EB resist with different degree of defects and cylinder structures of PS-b-PMMA, 300
    SEM image of spherical and cylindrical microdomain by PS-b-PMMA, 296
  microphase separation in bulk and thin films, 288–91
    depictions of various architectures, 288
    experimental phase diagram of polysterene-block-polyisoprene diblock copolymer, 290
    orientation of lamellar structures under various interfacial affinities in thin films, 291
  self-assembly, 361
Boersch effect, 126
borosilicate crown glass (BSC), 162
bottom-up fabrication, 358–64
  miscellaneous techniques, 361–4
    SEM images of self-assembly of BCP, 363
    top and side view of an AAO template, 362
  nanomaterials from vapour deposition, 360–1
  solution-based synthesis of colloidal nanomaterials, 358–60
    TEM images of solution-synthesised nanoparticles and nanorods, 359
Bragg reflector, 47
Bragg wavelength, 47
brightness, 99
Broglie electron, 96
buffer gas, 56
bulk materials, 365–6
bulk micromachining, 430
bulk models, 444–5
bus waveguide distance, 380

C-shaped aperture combined with nanotip (CAN-Tip), 391–2
calixarene, 262–3
  chemical structure, 262
carbon nanotube (CNT)
  challenges, 412–13
  device integration, 412
Casino Monte Carlo simulation program, 105
catadioptric designs, 14
catoptric designs, 14
CD-AFM, 507–8
  principle of 3D AFM measurement, 508
CD-SEM, 506–7
  flowchart of measurement and characterisation of LER/LWR, 506
cell, 87
channelling, 134
character projection, 87
charge coupled device (CCD), 303
Charlesby-Pinner relationship, 251
'chemical freeze' process, 228
chemical heteroepitaxy, 297
chemical vapour deposition (CVD), 302, 360, 404
chemically amplified photoresists (CAR), 453–4
chemiluminescence, 435
chrome lift-off method, 174
chrome on glass (COG), 159
chromium nitride (CrN), 64
coefficient of thermal expansion (CTE), 522

coherence, 450–1
collector, 45
colloidal nanomaterials, 358–60
coma, 20
combined-nanoimprint-and-photolithography (CNP), 356–7
compact models, 445, 447–8
  convolution kernel-based modelling, 448
complementary metal-oxide semiconductor (CMOS), 319
computer aided design (CAD) system, 306
constant threshold model, 455
contact methods, 547
continuous flow microfluidics, 425–7
  schematic illustration of laminar flow concept in microfluidic channel, 426
  schematic of capillary effect between capillary and ambient pressures, 426
contrast, 199–200
contrast enhancement layers (CEL), 231
convolutions, 456
Cooperative Research and Development Agreement (CRADA), 49
core-shell structures, 360
Cotton-Mouton effects, 394
Coulomb blockade, 409, 410
Coulomb scattering, 94
Cramer–Rao bound, 478
critical dimension (CD), 83, 317, 504

Dai Nippon Printing (DNP), 334
debris, 56
deep silicon reactive ion etching, 430
depth of focus (DOF), 450
development time, 520
diazonaphthoquinone photoactive compound (DNQ-PAC), 201
diblock copolymer, 290–1
die-to-database inspection, 173
die-to-die inspection, 172
dielectric method, 429
digital pattern generator (DPG), 90
dioptric designs, 14
dip-pen lithography, 354–5
dip-pen nanolithography (DPN), 267
direct printing process, 431
direct write electron beam lithography exposure systems (DWEBLS), 234
directed assembly
  nanolithography, 287–311
    block copolymers in lithography, 287–94
    directed self-assembly of block copolymers, 294–302
    programmable three-dimensional lithography, 302–10
directed self-assembly, 294–302
discharge produced plasma (DPP), 53, 54
DNA hybridisation reactions, 436
double exposure lithography (DEL), 230–1
double patterning, 39, 43
double patterning lithography (DPL), 227–30
droplet microfluidics, 427–9
  schematic of droplet generation by T-junction consisting of rectangular channels, 428
  schematic representation of electrowetting on dielectric method for droplet actuation, 429
  schematic representation of electrowetting principle, 428
dry process, 292
3-D DXF format, 306
dynamic random access memory device (DRAM), 194

e-beam, 167
  imaging, 324–5
e-beam-induced deposition (EBID), 107
e-beam lithography, 298–9, 380, 457–60
  compact model, 459–60
  point spread function, 459
elastic recoil detection analysis (ERDA), 303

electrochemistry, 435
electrohydrodynamic sources, 120
electrokinetics flow mechanism, 427
electromagnetic field (EMF), 166
electromagnetic waves, 377
electron beam chemical-vapour-deposition (EB-CVD), 307–10
electron beam lithography, 80–111, 352–3, 432
  distributed systems, 100–3
    distributed fixed-axis system configuration, 102
    distributed system with multiple arrayed electron columns, 96
    ELIPS system concept, 102
  electron-beam patterning of photomasks for optical lithography, 107–10
    critical dimensional specification for photomasks, 109
    mask data volume, 110
  trade-off between resolution and throughput, 96–100
    resolution as function of beam semi-angle, 96
  typical positive-tone resist process, 81
  ultimate lithographic resolution, 104–7
    Monte Carlo simulation of individual electron trajectories in bulk silicon, 106
    ultimate lithographic resolution using EBID technique, 107
  using pixel parallelism to address the throughput bottleneck, 84–96
    exposure flashes for Gaussian beam system, 86
    Gaussian beam system configuration, 84
    MAPPER system configuration, 88
    multiple-column cell system concept, 89
    multiple shaped beam concept, 89
    PMLP system configuration, 88
    PREVAIL system concept, 91
    reflective electron beam lithography system concept, 90
    SCALPEL system concept, 91
    variable-shaped beam system configuration, 86
    writing speed in units of area swept out per unit time, 93
electron lithography image projection system (ELIPS), 101
electronic structure, 364–6
electroplating, 430
electrostatic chucks (ESC), 64
electrowetting, 428–9
embossing, 431
energy-beam-based lithography, 351–3
  electron-beam lithography, 352–3
  holographic lithography, 352–3
  photolithography, 352–3
energy dispersive X-ray spectroscopy (EDX), 343
Engineering Test Stand (ETS), 49
environmentally stable chemically amplified photoresist (ESCAP), 231
enzymatic assays, 436
etching resistance, 198
étendue, 55
exposure dose, 519
extraordinary optical transmission
  subwavelength apertures, 382–4
    enhancement and schematic of groove arrays and SEM image of fabricated structure, 384
    various antennas with gap sizes down to -10 nm and photodetector of dipole antenna, 386
extreme ultraviolet (EUV) lithography, 42–76, 460–2
  basics, 45–6
    schematic of an EUV multilayer mask in cross section, 46
  commercial phase, 51–2
    specifications for ASML's ADT, NXE:3100 and NXE:3300 platforms, 52
  development phase, 48–51
    chemically amplified EUV photoresist from Intel MET and SEMATECH LBNL MET, 51
    EUV limited liability company, 49–50

- national EUV lithography program, 49
- the consortia, 50–1
- EUV mask defects, 65–9
  - defect repair, 67–9
  - EUV mask blank defects, 67
  - opaque ad clear defects and their repair on EUV mask, 68
  - repaired absorber layer defect, 68
  - types of defects on EUV masks, 65
  - typical particle and pit defect on LTEM substrate, 66
- EUV masks, 62–70
  - EUV Pod from SEMI E152 specification, 69
  - fabrication, 63–5
  - handling, 69–70
  - IBD tool, 64
  - schematic of EUV mask fabrication process, 63
- EUV optics, 58–62
  - Mo/Si multilayer film stack, 60
  - multilayer deposition, 59–60
  - optical train, 61
  - optics lifetime, 61–2
  - PSD in arbitrary units of surface roughness on SEMATECH0.3 NA MET optic at LBNL, 59
  - six-mirror EUV projection optic design, 62
  - surface roughness, 58–9
- EUV resists, 70–3
  - approaches and progress, 73
  - challenges, 71–3
  - photolithography pattern transfer process, 70
  - requirements, 71
  - resolution, 71
- EUV sources, 52–8
  - approaches, 52–4
  - collector, 55–6
  - configurations for LPP and DPP EUV sources, 54
  - debris mitigation, 56–7
  - source fuels, 54–5
  - source scaling, 57–8
- future trends, 75–6
  - benefits, 75–6
  - limits, 76
- integration and implementation challenges, 73–5
  - out-of-band radiation, 73–4
  - system contamination, 74
  - system weight and power, 74–5
- key subsystems, 45
- research phase, 46–8
  - optical performance, resist performance in PMMA and safety considerations of candidate materials for multilayer, 48
- wavelength change, 43–5

extreme ultraviolet lithography (EUVL), 40, 317

fabrication methodologies, 429–33
- conventional, 430–2
  - schematic diagram of micromoulding process involving elastomers, 431
  - schematic representation of anodic bonding and microfluidic device fabrication, 430
  - schematic representation of hot embossing process, 432
- non-conventional, 432–3

Fabry-Pérot interferometers, 395
Faraday effects, 394
Fast Fourier Transform (FFT), 298
feature placement, 541
fibre Bragg gratings, 381
field effect transistor (FET), 409
field-programmable gate array (FPGA), 190
fin field effect transistor (FinFET), 523
finite-difference frequency-domain (FDFD) method, 377–8
finite-difference time-domain (FDTD) method, 377–8
finite integration technique (FIT), 377–8
flare, 21
Flory-Huggins interaction parameter, 288
fluid filling, 338–9
focal-spot writing, 182
focus, 518

focused-ion-beam chemical-vapour-deposition (FIB-CVD), 302–7
focused-ion-beam milling, 432
focused ion beams (FIB)
  adumbrated history, 118–19
  charged particle optics, 125–8
    four kinds of ion sources, and some of their parameters, 126
    generic two-lens optical column used with LMIS-based sources, 127
    shape of ion beam, 128
  deposition, 145–6
    Ga LMIS induced deposition, 146
    precursor gases used for depositing metals and insulators, 146
    water enhancement and retardation factor for different materials, 145
  future trends, 153–5
  imaging, 147–52
    array of nano-pillars created by GFIS He beam induced chemistry, 150
    cell body exposed to a virus, 151
    difference in contrast mechanisms available, 151
    layered sample designed to show the relative elemental contrast of several common materials, 152
    parameters and associated values for estimating secondary-electron yields of some insulators, 148
    same area showing contrast differences between electrons and Li ions, 152
    sputtering can limit the resolution available in an image, 149
    test pattern of conducting lines and pads produced by GFIS He ion beam, 150
  ion-matter interactions, 128–38
    channelling effect, 136
    channelling effect in Al by Ga, 136
    comparison of ranges, straggles, sputter yields and energy transferred to Si target atoms, 134
    density, sputter rate and sputter yield, 133
    deposition from a gas, 138
    depth of damage layer for Ga^+ into Si, 137
    electronic, nuclear and total stopping power of 30 keV Ga^+ across the periodic table, 132
    gas enhanced etching, 137
    geometry of range, projected range and transverse straggling distance for ion penetrating into the sample, 130
    implantation, sputtering, generation of secondary electrons and ions, 130
    nuclear and electronic stopping powers for He and Ga impinging into Si as function of energy, 131
    tabulation of TRIM calculations for Ga^+ incident on Si at normal incidence, 134
    TRIM calculations, showing the effect of incident energy, 135
    two-lens mass separated column, 129
  milling, 138–45
    array of 4 nm diameter nanopores drilled through a 30 nm thick silicon nitride membrane, 144
    'C'-shaped apertures milled in aluminium film, 144
    classes of FIB milling, 139
    crossover point between LMIS and PFIB, 141
    dual beam geometry used to 'slice-and-view' sample, 142
    graph comparing the d_{50} beam size diameter of Ga and Xe, 140
    identically shaped micro-fluidics channels directly milled in Si, 140
    penultimate milling step in creating a TEM sample, 139
    selective etching, 143
    three-dimensional reconstruction of dual beam geometry, 142

    - water-enhanced etching to remove PMMA resist, 143
  - nano machining and imaging, 116–55
    - schematic of major components of modern FIB system, 117
  - sources of ions, 119–25
    - AuGe LMIS emitter, 122
    - cold atom ion source (CAIS), 123–5
    - d_{50} spot size is the diameter of the beam that contains 50% of its current, 121
    - gas field ionisation sources, 122–3
    - inductively coupled plasma ion source (PFIB), 123
    - liquid metal ion sources (LMIS), 120–2
    - LMIS source showing the needle, reservoir, heater and base, 122
    - MOT-based CAIS system, 125
    - schematic of an inductively coupled ion source, 124
    - schematic of He GFIS source, 123
  - spectroscopy, 153, 154, 155
    - mass spectrum taken with Ga^+ beam, 154
    - SIMS elemental maps of the end region of the X chromosome of the fruit fly *D. melanogaster,* 155
- foil trap, 56
- Fourier transform (FT) analysis, 511
- Fresnel zone plates, 183
- fused quartz, 164
- fused silica, 164

- gap independence, 487–8
  - schematic diagram of beams diffracted in the transverse plane, 487
- gas field ionisation sources (GFIS), 117
- gate length roughness (GLR), 524
- Gaussian line shape analysis, 303
- glycidyl methacrylate-co-3-chlorostyrene (GMC), 252
- graded coatings, 61
- graphical-user interface (GUI), 191
- graphoepitaxy, 295
- grating formula, 7
- grating mask, 31
- gray-level scaling, 183
- grayscale writing, 87

- heat-assisted magnetic recording (HAMR), 392
- height-height correlation function (HHCF), 510, 512
- heterostructures nanowires, 413–15
- hexafluoroisopropyl alcohol (HFIPA), 222
- hexamethyldisilizane (HMDS), 167
- high-density semiconductor circuits, 317
- high index materials, 44
- high spatial frequency response (HSFR), 59
- Hoechst AG, 215
- holographic lithography, 352–3
- Hopkins diffraction model, 452–3
- hot embossing, 432
- hybrid-mask-mode (HMM) technique, 357
- hybrid nanoelectronics, 413–17
  - heterostructures nanowires building blocks, 413–15
    - schematic nanowire growth for both axial and radial heterostructures, 414
  - nanowire arrays the circuit regime, 415–17
    - schematic of nanowire crossbar configuration and SEM nanowire crossbar, 416
- hydrogen silsesquioxane (HSQ), 263–4
- Hypertext Transfer Protocol (HTTP), 191

- illumination optics, 45
- image analysis, 477–9
  - block diagram of basic phase analysis algorithm, 477
- image log slope (ILS), 11, 518
- image modulation, 11
- image placement error (IPE), 20

image projection, 181–2
imaging system, 475
immersion lithography, 36–8, 44
  comparison of immersion microscope setup in magnification mode and reduction immersion lithography setup, 37
  rays traced through the lithography stacks for dry lithography scenario and immersion lithography scenario, 38
imprint lithography, 320
imprint mask, 333–5
imprint tool, 331–3
'in-band' radiation, 55
incident power, 380
innate material roughness *see* resist roughness
integrated circuit (IC), 33, 430, 445
integrated phase reference, 492–4
  alignment marks and interference fringes, 493
  schematic of unambiguous phase alignment, 494
interconnect technology, 437
interferometers, 468–70
  alignment scheme of Flanders, Smith, and Austin, 469
interferometric lithography, 432
interferometric moiré alignment technique (i-MAT), 337
interferometric-spatial-phase imaging (ISPI), 467
  backside, 497–500
    images of interference fringe sets, 499
    mark design, 498
    schematic of a region of a backside ISPI mark, 500
  fundamentals, 472–5
    alignment information in spatial-phase discontinuity of interference patterns, 474
    schematic diagram, 474
  performance, 494–6
    alignment deactivity, 497
    sample of fringe pairs, 496
  photograph of microscope, 475
intermediate state two-photon (ISTP) materials, 231
International Risk Governance Council, 402
International Technology Roadmap for Semiconductors (ITRS), 2, 108, 178
intra-transistor roughness, 524
ion beam deposition (IBD), 64

jet and flash imprint lithography (J-FIL), 331–9
  imprint mask, 333–5
    full field patterned media template and imprinted bit patterns with 25 nm pitch, 335
    mask replication process imprints from master mask and pattern transferred on replica mask, 336
    photograph of J-FIL mask, 334
    SEM image of imprint mask for half pitches, 335
  imprint tool, 331–3
    optical projection stepper vs J-FIL stepper, 332
  performance results, 335–9
    field by field time budget for throughput of 20 wph, 339
    non-fill defects in 26 mm x 33 mm field and transition regions, 339
    overlay on 300 nm wafer using J-FIL tool, 338
    RLT uniformity across a 200 nm wafer, 337

Kelvin-probe, 407
KLA-T 2132, 343
KLA-T SP1 blank wafer inspection system, 343
krypton fluoride (KrF) laser, 203

Lagrange invariant, 16
laser-assisted direct imprint (LADI), 328–9
laser-induced fluorescence, 435
laser produced plasma (LPP), 53, 54
laser tweezers, 547
lateral force microscopy (LFM), 406–7

lens aberration, 17
light absorption, 367–8
light emission, 367–8
light oxygen descum process, 168
line edge/line width roughness (LER/LWR), 503, 512–13
  characterisation, 508–13
    correlaton functions, 510–11
    detected edges of a typical resist line, 509
    experimental measurements and fractal self-affinity, 513
    power spectrum, 511–12
    σ(L) curve, 512
line edge roughness (LER), 503–5
  effects on transistor performance, 525–9
    experimental results, 528–9
    modelling and simulation, 525–8
    random LER/LWR, 529
    schematic of effects of roughness on leakage currents, 529
  ideal shape of a line structure on a substrate, 504
  motivations for study, 505
  origins and mitigation, 530
  SEM image of a structure with lines and spaces, 505
linear zone plate (LZP), 467–8
  alignment scheme on a mask and narrow, orthogonal diffraction grating in a substrate, 468
liquid metal ion sources (LMIS), 116
liquid phase epitaxy (LPE), 403–4
liquid propulsion method, 426
litho-etch, litho-etch (LELE) process, 227–8
litho-freeze-litho-etch (LFLE) process, 228
litho-litho-etch (LLE) process, 228
lithographie galvanoformung abformung (LIGA), 430
lithography, 43, 351, 430
Littrow-angle, 490–2
  diffraction from a grating at various angles of incidence, 491
  schematic of drawings of diffracted beams from grating and checkerboard, 492
localised surface plasmon (LSP), 369–70
low dimensional devices
  nanodevices fabrication, prospects and applications, 399–417
    bottom-up hybrid nanoelectronics, 413–17
    future trends, 417
    motivation, 401–3
    nanofabrication creating building blocks for devices, 403–8
  prospects, 408–13
    dots to devices and 0th dimension to single-electron transistors, 408–11
    nano(tubes)wires, 411–13
low spatial frequency response (LSFR), 59
low thermal expansion material (LTEM), 58, 163
LumArray, Inc, 189–91

magnetic fields, 57
magnetic force microscopy (MFM), 406–7
magnification actuator system, 337–8
mask
  data preparation and processing, 159
  defects and photomasks, 172–7
    mask repair, 173–5
    patterned mask inspection, 172–3
    pellicles, 175–7
  for micro- and nanolithography, 158–77
  lithographic wavelengths and light sources, 162
  materials, 162–7
    requirements for absorber films, 165–7
    requirements for fused silica mask substrates, 164–5
    substrate materials, 162–4
    substrate materials and their most relevant properties, 163
  metrology, 168–72
    CDSEM images, 169

mask (*cont.*)
    critical dimension metrology, 169–70
    image placement metrology, 170–1
    phase and transmission metrology, 171–2
    schematic diagram showing optical signal from clear feature on a mask, 170
  process, 166, 167–8
    flow from pattern formation through pellicle mount, 166
  summary table of mask types, 160
  types, 159–60
  use, 161–2
mask level, 108
mask lifetime, 334–5
mask roughness, 515, 518
mask-substrate alignment
  implementation, 475–9
    broadband illumination, 476–7
    image analysis, 477–9
    imaging system, 475
    schematic of multi-line, simulated broadband light source, 476
  interferometric moiré fringes, 466–501
    background of alignment method, 467–72
    backside ISPI, 496–500
    characteristic formation, 479–94
    fundamentals of interferometric spatial-phase imaging, 472–5
    future trends, 500–1
    implementation, 475–9
    performance of ISPI, 494–6
maskless-lithography systems, 179
maskless photolithography, 179–91
  commercialisation of ZPAL by LumArray, Inc, 189–91
    multilevel alignment, 190–1
    photograph of LumArray's ZP-150 system, 190
    software, 191
    specifications of LumArray's ZP-150, 189
    throughput, 190
  extending the resolution of ZPAL, 187–9
    schematic of AMOL, 188
  focal-spot writing with an array of lenses, 183
  generic forms: image projection and focal-spot writing, 181–2
  mask-based and maskless lithography, 179
  printing press and desktop printer, 179–80
  proximity-effect correction, 186–7
    optical micrographs of pattern as written without proximity-effect correction, 187
  use photon as opposed to charged particles, 180–1
    benefits of photons, 181
    instability of charged-particle systems, 180
  via focal-spot writing, 182–3
  zone-plate-array lithography, 183–6
    patterns written in photoresist using ZPAL, 186
    results in initial research at Massachusetts Institute if Technology, 185–6
    scan stage rather than beams, 185
    schematic depiction of ZPAL, 184
    use of diffractive-optical elements, 184–5
mass limited targets, 57
materials, 438
mathematical models, 444–62
  difficulties in modelling, 455–7
    model calibration, 455–6
    model runtime and accuracy, 456–7
  extreme ultraviolet/electron beam lithography modelling, 457–62
    e-beam lithography, 457–60
    EUV lithography, 460–2
  model critical dimension extraction, 454–5
  modelling technique properties, 462
  optical lithography modelling, 445–9
  optical system in optical lithography, 449–53
  photoresist model, 453–4
Maxwell's equation, 377
mean free path, 104
mechanical forming, 351–2, 355–8

memory effect, 231
metal-dielectric-metal (MDM) structure, 393
metal organic chemical vapour deposition (MOCVD), 404
metal-oxide-semiconductor field-Effect Transistor (MOSFET), 523
  devices, 367
metal–insulator–metal (MIM), 539, 540
metallic nanostructures
  optical properties, 368–70
    schematics of surface plasmon polaritons excited by light, 369
micro-stereolithography, 432
micro-wine-glass, 304
microcontact printing *see* soft lithography
microdomain structures, 289
microelectronic processing, 350–1
microfluidic channel geometry, 427
microfluidic solution, 547
microfluidics, 424–41
  applications, 434–6
    biological and biomedical, 436
    chemical and environmental, 434–5
    emerging applications, 436
    energy, 435–6
  current trends, 425–9
    continuous flow microfluidics, 425–7
    droplet microfluidics, 427–9
  further information sources and advice, 440–1
    microfluidic companies, 441
    research groups in academia and national laboratories, 440–1
  future trends, 436–9
    commercialisation, 439
    technology challenges, 437–8
  present state of technology, 429–34
    fabrication methodologies, 429–33
    interconnect approaches, 433–4
microlithography, 158
micromachining technology, 175
micromoulding, 430
microphase separation, 288–91
microprocessor (MPU), 194
microring resonator devices, 326–7
Microsonic Laser Systems AB, 181
mid spatial frequency response (MSFR), 59
Millipede project, 357–8
model-based OPC, 447
model critical dimension extraction, 454–5
  constant threshold model, 455
  variable threshold model, 455
model form, 457
modulation transfer function (MTF), 11
moiré
  characteristic fringer formation, 479–94
    gap independence, 487–8
    integrated phase reference, 492–4
    Littrow-angle viewing, 490–2
    low NA, long working distance optics, 488–90
    refractive-index independence, 483–7
    schematic of a conventional brightfield microscope, 489
    wavelength independence, 479–83
  patterns, 470–2
    alignment marks used in the OAI alignment scheme, 473
    schematic of circular moiré patterns, 471
    schematic of moiré cross produced by LZP arrangement, 472
molecular beam epitaxy (MBE), 404
molecular models, 444
molecular nanosystems, 403
molybdenum oxide and silicon dioxide APSM (Mo-Si APSM), 33
Monte Carlo simulation, 99
Moore's Law, 42, 76, 402, 538
moulding, 430
multiple aperture pixel-by-pixel enhancement of resolution (MAPPER), 88
multiple-column cell projection (MCC), 89
multiple patterning, 43
  optical lithography, 38–9
multiple shaped beam (MSB), 89
mutual phase referencing, 492

NAND Flash memory, 319
nanodevices
  fabrication, prospects for low dimensional devices and applications, 399–417
    bottom-up hybrid nanoelectronics, 413–17
    future trends, 417
    nanofabrication creating building blocks for devices, 403–8
  motivation, 401–3
    representation development suggested by Renn and Roco to International Risk Governance Council, 402
nanofabrication
  creating building blocks for devices, 403–8
  materials characterisation structure, 406–8
    schematic representation of main components used to scan sample using AFM, 407
  materials growth source, 403–6
    schematic of MBE growth chamber, 405
    traditional plot of bandgap energy and wavelength vs lattice constant, 406
nanoimprint lithography (NIL), 355–8
  alternative thermal imprint processes, 327–30
  defectivity, 342–4
    non-fill defect, plug defect and transition region defect, 343
    results for an extended imprint run, 344
  history, 315–17
    Gutenberg press and page from 42 line Bible, 316
  jet and flash imprint lithography, 331–9
  overview, 320
    master mask of deposited2.4 nm carbon nanotube and imprint of carbon nanotube, 320
  rising cost, 317–20
    cost of ownership at 22 nm for different lithographic processes, 318
    cost of ownership for several markets requiring high resolution patterning, 319
    transistor density and tool cost as function of time, 318
  roll to roll imprint lithography, 340–2
  soft lithography, 321–2
  thermal imprint lithography, 322–7
  ultraviolet (UV) overview, 330–1
    schematic illustration of step and flash imprint lithography, 331
nanoindentation, 357–8
nanolithography
  chemistry and processing of resists, 194–271
    chemically amplified resist process considerations, 215–17
    chemically amplified resists for 193 nm lithography, 217–31
    resists for electron beam lithography, 234–66
    resists for extreme ultraviolet lithography, 231–4
    resists for optical lithography, 200–15
    resists for selected forward looking lithographic technologies, 266–71
  new applications and emerging technologies, 538–48
    geometry control of tunnel junctions, 543–7
    high-resolution patterning to new device structures, 541–3
    quantum dot placement problem, 547–8
  sidewall roughness, 503–31
    experimental results, 515, 518–23
    impact on device performance, 523–9
    metrology and characterisation, 505–13
    modelling and simulation, 513–15

nanophotonics
  devices for manipulating light at nanoscale, 376–96
    applications in photovoltaics, 395
    extraordinary optical transmission through subwavelength apertures, 382–4
    near-field optical microscopy, 390–2
    nonlinear processes enhancement, 393–5
    optical nanoantennas, 384–7
    photonic crystals, 377–9
    plasmonic focusing, 387–90
    plasmonic waveguides, 392–3
    ring resonators, 379–82
NanoRuler, 470
nanoscale
  electron source generation, 392
  nanophotonics devices for manipulating light, 376–96
  precision, 547
nanoscale lithography, 291
  techniques, 351–2
nanostructures
  electrical and optoelectronic properties, 364–8
    electronic transport, 366–7
    optoelectronic properties, 367–8
    quantum confinement, electronic structure and density of states, 364–6
  fabrication and applications, 348–71
    applications, 370–1
    characterisation, 350
    methods to create nanostructures and bottom-up fabrication, 358–64
    methods to create nanostructures and top-down fabrication, 350–8
  properties, 364–70
    optical properties of metallic nanostructures, 368–70
nanosystems, 403
nanotechnology, 402
nano(tubes)wires
  1-D to circuits, 411–13
    carbon nanotube (CNT) device integration, 412
near- infrared (IR) wavelength, 368–9
near-field optical microscopy, 390–2
  aperture probe NSOM, 390
  apertureless probe NSOM, 390–1
  CAN-tip probe NSOM, 391–2
    response when scanning across 3 Cr nanodisk and close-up of first disk scan, 392
    schematic design and 35° angled view SEM images focused ion beam milling, 391
near-field scanning optical microscopy (NSOM), 350, 390
near-field tip lithography, 354–5
next-generation lithography (NGL), 76, 317
Nikon, 51
non-contact methods, 547
nonlinear photonic crystal microresonators, 394–5
normalised image log slope (NILS), 11
novolac resin, 201
numerical aperture, 488–90
  lens geometries, 489
numerical simulations, 515
  analytical and numerical models of LER, 516–17

1-octadecanethiol (ODT), 267
off-axis illumination (OAI), 27
on-axis interferometric (OAI) alignment, 471
optical invariant, 36
optical lithography, 370
  modelling, 445–9
    compact models, 447–8
    flow diagram of lithography model, 449
    image formation modelling, 449
    optical proximity correction, 445–7
  optical system, 449–53
optical mask, 161
optical nanoantennas, 384–7
optical path difference (OPD), 12

optical projection lithography, 1–40
  fundamentals of optical lithography, 5–8
    basic components of projection lithography system, 6
    optical elements of projection imaging system, 6
  image evaluation, 8–13
    aerial image of space feature or mask opening, 10
    comparison of wavefronts corresponding to two focus positions, 13
    depth of focus, 11–13
    image quality, 10–11
    image resolution, 8–10
    projection imaging with coherent illumination, 9
    projection imaging with partially coherent illumination, 9
  immersion lithography, 36–8
    comparison of immersion microscope setup in magnification mode and reductin immersion lithography setup, 37
    rays traced for dry lithography and immersion lithography scenario, 38
  lithography in the deep-UV, 23–7
    catadioptric projection objection lens using beam splitting cube, 27
    chromatic aberration as change in focus for different wavelengths for refractive lens, 25
    fractional wavelength shrink from one lithography generation to the next, 24
    reflective design approaches leading to partially reflective alternatives, 26
    use of cemented doublets in refractive lens allow for the correction of chromatic aberration, 26
  lithography technology and trends, 1–5
    critical dimension and pitch designations defined for semiconductor geometry, 2
    methods used for 40 years of IC generations, 3
    microprocessor design evolution, 4
    wavelength scaling trends, 4
  modified illumination, 33–6
    customised illumination and source-mask optimisation, 35–6
    off-axis illumination, 33–5
  multiple patterning optical lithography, 38–9
  projection lithography systems, 13–22
    aberrations, flare and their influence, 17–21
    alignment and overlay, 21–2
    comparative plot of information content of various lithography systems measured as Lagrange invariant, 17
    comparison of maximum field sizes for stepper and scanner lithography systems, 14
    exposure sequencing for step-and-repeat lithography system and step-and-scan lithography system, 14
    hyper-NA optical projection systems for immersion lithography, 18
    mostly reflective optical design of full wafer scanner, 15
    numerical aperture and field size trends, 16–17
    projection lithography optics, 14–16
    refractive design example of reduction projection lithography lens, 15
    wavefront variation in objection lens pupil in the presence of aberration, 19
    Zernike aberration terms for primary aberrations, 20
    Zernike polynomial expansion representations of individual aberration types, 20
  resolution enhancement technology, 27–36
    aerial image plots for a mask with 100 nm lines of varying duty ratio, 30

      aerial image plots for lines of two duty ratios, 29
      comparison of binary mask and alternating PSM, 32
      comparison of capture of diffraction order energy for on-axis illumination and dipole OAI, 35
      distribution of diffraction energy with OAI and its capture by the objective lens, 34
      mask correction approaches and optical proximity correction, 28–31
      OPC using small serifs, 31
      OPE for lines of various pitch values through-focus, 29
      phase shift masking, 31–3
      several of the choices for OAI in projection lithography system, 34
   wavelengths for optical lithography, 22–3
      spectral distribution of emission lines of mercury vapour lamp, 23
optical proximity correction (OPC), 27, 445
   schematic representation of concept, 446
optical proximity error (OPE), 28
optical resolution, 488
optical ring resonators, 381
optical system
   optical lithography, 449–53
      Abbe's model, 451–2
      Hopkins diffraction model, 452–3
      partially coherent imaging system, 449–51
optical threshold layers (OTL), 231
optoelectronic devices, 370–1
optofluidics, 436
order-disorder transition (ODT), 289

Parseval's theorem, 512
partial coherence, 450–1
passive nanostructures, 403
pattern area, 456–7, 457
pattern generators, 179
pattern sampling area, 457
pellicle, 63, 176
periodic grading coupling, 544
phase reference, 474
phase shift masking (PSM), 27
photo-acid generator (PAG), 72
photo-polymerisation, 332
photoacid generator (PAG), 453, 521
photolithography, 298–9, 352–3
   stepper, 332–3
photomask, 159
photon-limited, 53
photonic band gaps (PBG), 377
photonic crystals, 309, 377–9
   2-D square with lattice constant, point defect, line defect and transmission spectrum, 379
   band gap, 377–8
   defects and cavities, 378–9
photonic nanostructures, 371
photoresist model, 453–4
photovoltaics, 395
physical vapour deposition, 360
physics behind phenomena, 438
pitch division, 39
pitch splitting, 39
Planck's constant, 367
plane wave method (PWM), 377–8
plasmonic focusing, 387–90
   schematics of adiabatic nanofocusing s-NSOM and SEM image of conical gold tip, 389
   schematics of metallic rod with taper angle, 388
   schematics of V-groove and wedge plasmonic waveguides, 388
plasmonic imaging, 350
plasmonic nanofocusing, 389–90
plasmonic waveguides, 392–3
point spread function (PSF), 106, 185, 452, 461
polarised illumination, 27
poly-(4-tert-butoxycarbonyloxystyrene) (PTBOC), 205
polydimethylsiloxane (PDMS), 268, 321
polyhydroxystyrene (PHS), 73
polymer pen nanolithography (PPL), 268
polymer resist effects, 520–1
polymethylmethacrylate (PMMA), 231, 323–4

post-exposure bake (PEB), 519
  diffusion, 454
power spectral density (PSD), 58, 520
pre-exposure data preparation (PEDP), 191
programmable three-dimensional lithography, 302–10
  EB-CVD, 307–10
    cubic photonic crystal constructed in prefabricated gap, 309
    high resolution TEM image proves wall roughness, 308
    Smith-Purcell electron optics, 310
  FIB-CVD, 302–7
    data flow of 3-D pattern-generating system, 306
    fabrication process illustration, 304
    Micro Leaning Tower of Pisa and microcoil illustration, 305
    Micro Starship Enterprise NCC-1701D, 3-D CAD model and SIM image, 307
    micro-wine glass on Si substrate and human hair, 305
    T-4 bacteriophage, 307
projection-electron lithography (PEL), 90
  tools, 235
projection lithography systems, 13–22
  aberrations, flare and their influence, 17–21
  alignment and overlay, 21–2
  comparative plot of information content of various lithography systems measured as Lagrange invariant, 17
  comparison of maximum field sizes for stepper and scanner lithography systems, 14
  exposure sequencing for step-and-repeat lithography system and step-and-scan lithography system, 14
  hyper-NA optical projection systems for immersion lithography, 18
  mostly reflective optical design of full wafer scanner, 15
  numerical aperture and field size trends, 16–17
  projection lithography optics, 14–16
  refractive design example of reduction projection lithography lens, 15
  wavefront variation in objection lens pupil in the presence of aberration, 19
  Zernike aberration terms for primary aberrations, 20
  Zernike polynomial expansion representations of individual aberration types, 20
projection maskless lithography patterning (PMLP), 87
projection reduction exposure with variable axis immersion lenses (PREVAIL), 90
Prolith, 445
propagating surface plasmon (PSP), 369–70
prototype three-dimensional models, 432
proximity control, 543
proximity effect, 83, 105
Purcell effect, 395

quantum confinement, 364–6, 539
  nanostructures vs bulk materials, 365
  schematic of electron confinement or delocalisation from atom to bulk materials, 365
quantum dot, 370–1, 404–5, 547–8
  experimental setup and control principle, 548
quantum mechanics theory, 409
quantum size effect, 409
quantum wells, 370–1, 404–5
quarter wave limit (QWL), 21
quencher, 521

Raman microscopy, 437
Raman spectra, 303
Raman spectroscopy, 370, 437
raster pattern, 85
Rayleigh criterion, 43
Rayleigh equation, 8, 27, 39
reactive gas cleaning, 57
rectenna, 542

reflective electron beam lithography (REBL), 90
refractive-index independence, 483–7
  immunity of ISPI to overlayers of resist, polysilicon, and aluminium, 486
  ISPI image of fringes from an alignment mark, 485
  sketch of diffracted beams inside resist, 484
replica moulding, 431
residual layer thickness (RLT), 335–6
resist, 80
  acid catalysed reactions, 208–15
    epoxy-containing resins used in negative acting CA resist formulations, 209
    generalised chemical structures for protected polyhydroxystyrene and polymetacrylic acid, 213
    high fidelity coded 0.18 μm lines and spaces in CA positive acting resist, 212
    negative acting chemically amplified resists, 208–11
    positive acting chemically amplified resists, 211–15
    schematic representation for radiation induced reaction mechanism occurring in PTBSS, 214
    three-component tert-butylcholate based CA resist formulation, 213
    two-step radiation induced chemical amplification reaction process, 215
  chemically amplified resist process considerations, 215–17
    patterning process flow diagram for CA resists requiring PEB step, 216
  chemically amplified resists for 193 nm lithography, 217–31
    acrylate-based monomers containing hexafluoroisopropanol group, 222
    chemical structure for the terpolymer present in IBM ‘version-1’ 193 nm resist, 219
    chemical structures for alicyclic polymers incorporated into 193 nm resists, 220
    diagrammatic representation of ‘Litho-chemical freeze litho etch,’ 228
    double exposure lithography, 230–1
    double patterning lithography, 227
    fluorine-containing copolymer materials functioning as top coats, 225
    immersion compatible/top coat free resists, 225–7
    LELE, LLE, LFLE and SADP process flow diagrams, 230
    monomers comprising the resist in a litho-thermal freeze litho-etch process’, 229
    resists for double patterning, double exposure 193 nm lithography strategies, 227–30
    silicon and fluorine-containing methacrylate-co and terpolymers used as additives in top coat free 193 nm immersion lithography, 226
    top coat protective films, 223–5
    two-component resist formulations and three-component resist formulations, 221
    UV absorption spectra of PHS and polyacrylate based polymers, 218
  chemically amplified resists for 248 nm lithography, 203–8
    acid generator chemistry, 206–8
    DUV photolysis and acid formation mechanism for 2-nitrobenzylesters, 208
    generalised DUV photolysis mechanism for triphenylsulfonium based ionic acid generating species, 207
    schematic representation of two-step chemical amplification process, 205

resist (*cont.*)
    UV-Visible absorption spectra of novolac and polyhydroxystyrene polymers, 204
  chemistry and processing for nanolithography, 194–271
    generic exposure response/contrast curves for positive and negative acting resists, 200
    influence of lithography exposure system on required sensitivity and performance, 198
    list of exposure technology and type of resist used to delineate a minimum resist feature size ranging from 0.10 to 1.0 μm, 197
    minimum DRAM and MPU device feature size vs the first year of IC production, 195
    process flow diagram for the resist pattern delineation process, 197
    process flow for subtractive and additive resist pattern transfer step, 199
    required resist sensitivity/process dose vs lithography exposure technology, 198
    resist design considerations, 196–8
    resist performance characterisation methodologies, 199–200
    spectral output of typical high pressure Hg arc lamp, 201
    stability of the resist in the pattern transfer environment, 198–9
  dip-pen/polymer pen nanolithography, 267–8
  electron beam lithography, 234–66
    cage and network forms of HSQ inorganic resist, 264
    calixarenes, 261–3
    change in molecular weight as measured by gel permeation chromatography vs e-beam exposure dose for PMMA positive acting resist, 238
    chlorinated styrene resists, 256
    coded 0.15 μm line and space patterns in a PMPS/novolac blended positive action electron beam resist, 245
    coded 15 nm dot array in calixarene resist patterned using direct write e-beam lithography system, 263
    coded 80 nm isolated trench patterns, 246
    components in a negative acting electron beam sensitive resist, 261
    Cr-on-glass photomask electron beam resist property vs performance requirements, 236
    dense array of coded 80 nm contact holes patterned in 750 nm thick positive acting CA resist, 249
    e-beam exposure response curves for GMC, 254
    e-beam induced reaction mechanism in poly(2-methyl-1-pentene-co-sulfur dioxide), 244
    electron beam induced degradation mechanism in PBS, 241
    electron beam induced reaction pathway in PGMA negative acting resist, 250
    electron beam induced reaction sequence occurring in the chlorinated aromatic negative acting resist, 255
    etch selectivity values of P(SI-MS), 258
    fraction of mass of the acid remaining vs heating time, 247
    hafnium and zirconium based resists, 265
    high aspect ratio sub-100 nm isolated lines imaged in NEB22A resist, 262
    hydrogen silsesquioxane, 264–5
    intrinsic property requirements of polymers as electron beam sensitive resists, 236–8
    lithographic property vs performance results for Sumitomo NEB 22A negative acting resist, 261

material properties of range (P(SI-CMS) formulations, 258
negative acting chemically amplified electron beam resists, 259–61
negative acting non-chemically amplified electron beam resists, 248–58
oligomeric resin component in SU-8 resist, 259
orthogonal coded 50 nm equal lines and spaces formed in bilayer resist stack, 264
oxygen plasma dry etch rate ratio, 257
plot of contrast value for e-beam sensitive PCMS resist, 256
plot of gel fraction vs absorbed dose for GMC copolymer, 252
plot of reciprocal number average molecular weight vs absorbed dose, 239
positive acting electron beam resists, 238–46
positive acting resists for high resolution and throughput electron beam lithography, 245–8
profiles of coded 300 nm line and space patterns, 246
property vs performance requirements for resists, 253
radiation induced reaction process in polymethylmethacrylate resist, 239
repeat unit for the 1-olefin sulfone copolymers, 241
repeat unit in poly(glycidyl methacrylate-co-3-chlorostyrene) negative acting electron beam resist, 252
repeat unit in P(Si-CMS), 257
repeat unit in the ZEP family of positive acting electron beam resist, 243
sequence of steps and conditions used in the P(SI-CMS) Cr dry etch optical mask pattern transfer process, 258
side chain/pendent groups on an acrylate backbone polymer, 250
step wise fundamental analysis approach to determine predominant e-beam induced reaction process, 237
sub-25 nm line patterns formed in 100 nm thick SU-8 film, 260
submicron vertical Cr line profiles, 257
submicron vertical pattern profiles in SPR 700 i-line resist, 243
two positive acting e-beam resists containing high electronegativity/electron affinity substituents, 240
undercut profile of a <1.0 coded μm Cr line formed using a wet etch pattern transfer process, 242
for extreme ultraviolet lithography, 231–4
  acid generating species incorporated into CA EUVL resists, 234
  chemical structure and chemical amplification reaction process in KRS resist, 232
  EUVL acid generating species, 233–
  hydroxystyrene-based polymer platform, 231–2
  ketal protected hydroxystyrene-based polymer platform, 232–3
  low EUV absorption cross-section resists, 233
  methacrylate-based polymer platform, 231
  polymer/resin component in ESCAP resist, 232
  silicon containing resists, HSQ and boron containing positive and negative acting EUVL resist polymers, 233
for nanoimprint lithography, 266–7
  select monomers, photoinitiator and crosslinking species incorporated into UV-NIL resists, 267

resist (*cont.*)
    thermal nanoimprint lithography (T-NIL), 266
    UV nanoimprint lithography (UV-NIL), 266
  for optical lithography, 200–15
    chemical structure representation of components in a standard photoresist formulation, 202
    DNQ-PAC photolysis reaction sequence, 203
    positive acting photo resists for g- and i-line lithography, 201–3
  for selected forward looking lithographic technologies, 265–8
  resolution limitations, 268–70
    diagrammatic representation of resist pattern collapse process, 269
    illustrates a model for the aerial image blurring processes in a CA resist, 270
    resist blur, 269–70
    resist pattern after completion of the wet development and drying steps, 269
    resist pattern collapse, 268–9
resist roughness, 521
resist-substrate interactions, 522
resolution, 541
resolution enhancement technology (RET), 27–36
  aerial image plots for a mask with 100 nm lines of varying duty ratio, 30
  aerial image plots for lines of two duty ratios, 29
  comparison of binary mask and alternating PSM, 32
  comparison of capture of diffraction order energy for on-axis illumination and dipole OAI, 35
  distribution of diffraction energy with OAI and its capture by the objective lens, 34
  mask correction approaches and optical proximity correction, 28–31
  modified illumination, 33–6
    customised illumination and source-mask optimisation, 35–6
    off-axis illumination, 33–5
  OPC using small serifs, 31
  OPE for lines of various pitch values through-focus, 29
  phase shift masking, 31–3
  several of the choices for OAI in projection lithography system, 34
reticle, 159
Reynolds number, 425
ring resonators, 379–82
  active optical devices, 381–2
  biochemical sensors, 382
  optical filters, 381
  schematic of four-port microring resonator and on-chip four-port, 380
RLS triangle, 520, 530
roll to roll imprint lithography, 340–2
  process sequence for HP's SAIL process, 340
  schematic roll to roll and roll to plate imprinter and tool image, 342
  web based imprint tool used to create flexible displays, 341
roller mask, 341–2
rotating foil trap, 57
roughness exponent, 510
rule-based OPC, 446–7
runtime, 456–7
Rutherford backscattering spectrometry (RBS), 303

sampling rates, 456
SCALPEL system, 246
scanning beam interference lithography (SBIL), 470
scanning-electron-beam lithography (SEBL), 180
scanning electron micrograph, 295–6
scanning electron microscope (SEM), 116, 329, 350, 384, 504
scanning ion microscope (SIM), 306–7
scanning probe lithography, 355

scanning probe microscopy (SPM), 406–7
scanning-tunneling microscope (STM), 350
scattering near-field scanning optical microscopy (s-NSOM), 389–90
SCattering with Angular Limitation Projection Electron Beam Lithography (SCALPEL), 90
scatterometry, 508
Schottky emitter, 85
second-harmonic generation (SHG), 394
secondary-ion mass spectrometry (SIMS), 116
secondary plasma, 57
SELETE, 50
self-aligned imprint lithography (SAIL), 340
Self-Aligned Spacer Double Patterning Process (SADP), 227, 229
self-assembled monolayer (SAM), 322
self-assembly
  anodised aluminium oxide (AAO), 361
  block-copolymer (BCP), 361
self-consistent field theory (SCFT), 289
SEMATECH, 50
SEMATECH LBNL MET, 70
SEMI E152, 69
Semiconductor Clean 1 (SC-1), 168
semiconductor mask fabrication, 333
Sentarus Lithography, 445
sharp electron beam, 355
shells, 56
Shipley, 260
short-pulse lasers, 173
shot noise, 72, 519
sidewall roughness
  experimental results, 515, 518–23
    LER during pattern transfer, 522–3
    mask roughness effects, 515, 518
    material effects, 520–2
    post-processing treatments for reducing LER, 522
    process effects, 518–20
  impact on device performance, 523–9
    impact of resist on mean and standard deviation, 527
    LER effects on transistor performance, 525–9
    LER/LWR,gate length roughness and CD uniformity, 524–5
    rms of GLR and CD variation vs the gate width or line length, 525
  metrology and characterisation, 505–13
    characterisation of LER/LWR, 508–13
    measurement techniques, 506–8
  modelling and simulation, 513–15
    analytical modelling, 514–15
    lithographic process steps, 514
    numerical simulations, 515
  nanolithography, 503–31
signal-to-noise ratio, 478
silicon substrate, 324–5
single-electron transistor (SET), 367, 409
  basics of operation, 409–11
    schematic illustration, 410
  challenges, 411
SMI2050 system, 302
SMI9200 system, 302
Snell's Law, 36, 483
soft lithography, 321–2, 431
  30 nm ring created with μCP and working field effect transistor, 323
  process schematic, 321
solar cells, 395
solution-based synthesis, 358–60
source-mask optimisation (SMO), 35–6
space charge, 126
spatial coherence, 5
spatial phase, 486
spectral purity filter (SPF), 74
spin coating, 291–2
spin-on techniques, 330–1
sputter yield, 131
stage-limited, 53
stamping, 540
standard aperture theory, 382–3
standardised components, 438
state density, 364–6

step-and-flash imprint lithography (S-FIL), 330–1
stimulated emission depletion (STED), 189
stray light, 21
stress birefringence, 165
subwavelength apertures, 382–4
Sumitomo Chemical, 261
surface-enhanced Raman scattering, 371, 394
surface-enhanced Raman spectroscopy, 435
surface micromachining, 430
surface modification techniques, 299
surface plasmon (SP) polaritons, 354

T-junction, 427
Taylor cone, 121
Teflon, 342
tert-butoxycarbonyloxy (TBOC), 205
tetramethylammonium hydroxide (TMAH), 168, 205
'thermal freeze' process, 229
thermal imprint lithography, 322–7
  features in mould and in imprinted resist, 325
  microring resonator mould and output spectra response with TM-polarised input, 328
  parametric variance of beam at wavelength of632.8 nm, 327
  schematic flow process, 324
  sub-10 nm holes formed using T-NIL, 326
thermal nanoimprint, 357
thin-film dielectric interference filters, 381
thin-mask approximation (TMA), 166
thin oxide film deposition, 325
three-dimensional (3D) microfluidic devices, 433
tip-based lithography, 351–2, 354–5
tip-enhanced Raman scattering (TERS), 392
Tokuyama Corp, 262
top-down fabrication, 350–8
  energy-beam-based lithography, 352–3
  mechanical forming and nanoimprint lithography, 355–8
    schematics of NIL, thermal and UV-curable nanoimprint, 356
  tip-based, near-field tip and dip-pen lithography, 354–5
    schematics of tip-based imaging and lithography system, 354
  top-down and bottom-up routes and their characteristic length scales, 351
total integrated scatter (TIS), 460
transmission electron microscope (TEM), 302–3, 350, 403
TRansport of Ions in Matter (TRIM), 134
transverse electric (TE) mode, 378
transverse magnetic (TM) mode, 378
trim mask, 32
tunneling, 539, 540
  application of high-resolution patterning to new device structures, 541–3
    basic patterns and location of tunnel junction diode, 542
  geometry control of junctions, 543–7
    asymmetric antenna fabrication, 543–4
    future of beam-based lithography, 543
    pattern placement problem, 544, 547
tunnelling, 366–7

ULE, 164
ultraviolet mask, 161
ultraviolet nanoimprint lithography (UV-NIL), 330–1
ultraviolet (UV), 353
  curable nanoimprint, 356–7
  imprint stepper, 332–3
  treatment, 292
usable depth of focus (UDOF), 12

van der Waals forces, 407–8
vapour phase deposition, 360–1
vapour phase epitaxy (VPE), 403–4
variable threshold model, 455

vector scanning, 85
very large scale integration (VLSI) designs, 413
Virtual National Laboratory (VNL), 49
visible wavelength, 368–9

wafer bonding, 430
wafer stage motion, 338
waveguide refractive index, 380
wavelength independence, 479–83
  schematic of diffraction from a substrate grating, 480
  schematic of double diffraction from different grating periods, 480
wet etch, 168, 322
Wien filter mass separator, 117

X-ray absorption, 303
X-ray lithography, 333
X-ray proximity lithography (XPL), 46
X-ray range, 353
xenon, 55

yttrium aluminium garnet (YAG) lasers, 173

Zernike polynomial description, 19
Zerodur, 164
zeroth dimension, 409
zone-plate-array lithography (ZPAL), 183–6
ZP-150A tool, 189–91